

University of Dundee

Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10

Forgione, Rosa Ester; Di Carluccio, Cristina; Guzmán-Caldentey, Juan; Gaglione, Rosa; Battista, Filomena; Chiodo, Fabrizio

Published in:
iScience

DOI:
[10.1016/j.isci.2020.101231](https://doi.org/10.1016/j.isci.2020.101231)

Publication date:
2020

Licence:
CC BY-NC-ND

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Forgione, R. E., Di Carluccio, C., Guzmán-Caldentey, J., Gaglione, R., Battista, F., Chiodo, F., Manabe, Y., Arciello, A., Del Vecchio, P., Fukase, K., Molinaro, A., Martín-Santamaría, S., Crocker, P. R., Marchetti, R., & Silipo, A. (2020). Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10. *iScience*, 23(6), 1-29. [101231]. <https://doi.org/10.1016/j.isci.2020.101231>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

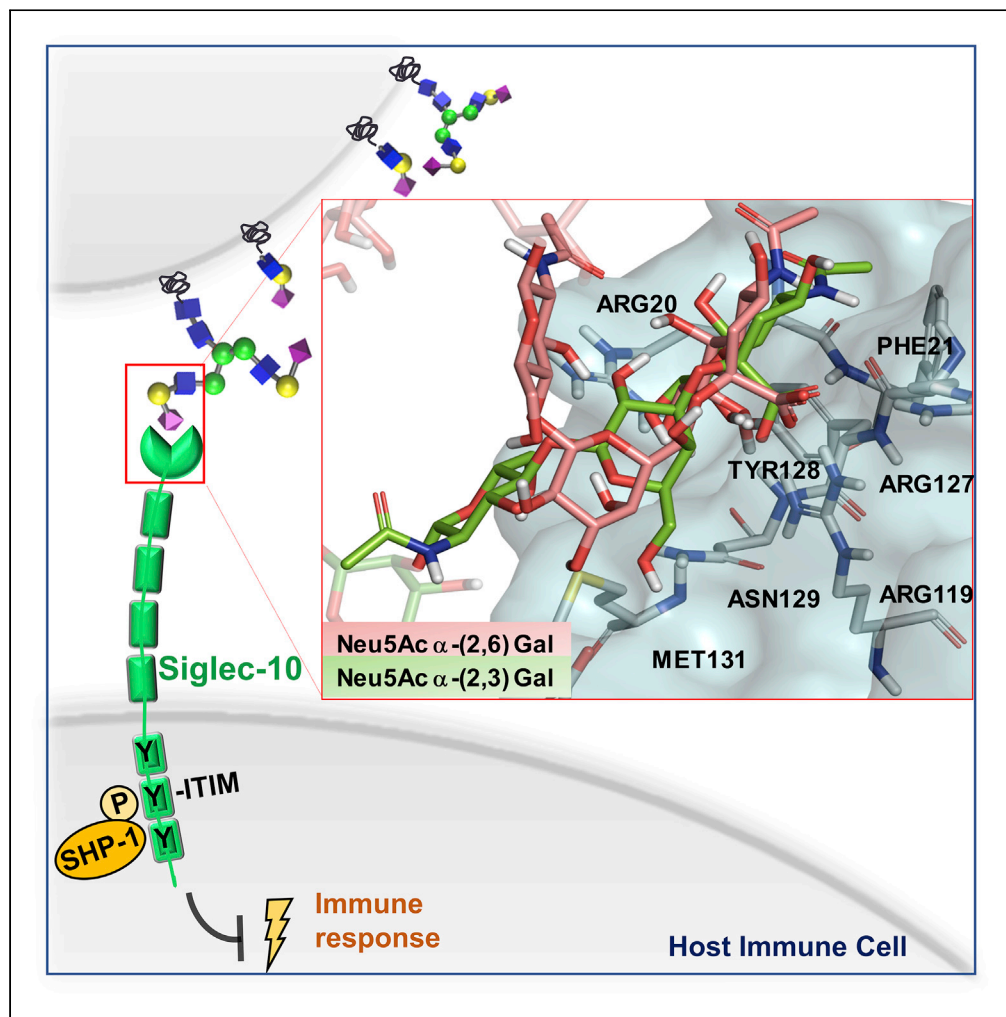
- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Article

Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10



Rosa Ester Forgiione, Cristina Di Carluccio, Juan Guzmán-Caldentey, ..., Paul R. Crocker, Roberta Marchetti, Alba Silipo

roberta.marchetti@unina.it (R.M.)
silipo@unina.it (A.S.)

HIGHLIGHTS

We unveiled the molecular basis of sialoglycans recognition by Siglec-10

The conformation of sialoglycans drives the interaction with the protein

Siglec-10 is able to recognize and bind complex N-glycans

Our outcomes may open the venue for the design and development of novel glycomimetics

Forgione et al., iScience 23, 101231
June 26, 2020 © 2020 The Author(s).
<https://doi.org/10.1016/j.isci.2020.101231>

Article

Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10

Rosa Ester Forgiione,¹ Cristina Di Carluccio,¹ Juan Guzmán-Caldentey,² Rosa Gaglione,¹ Filomena Battista,¹ Fabrizio Chiodo,³ Yoshiyuki Manabe,⁴ Angela Arciello,¹ Pompea Del Vecchio,¹ Koichi Fukase,⁴ Antonio Molinaro,^{1,4} Sonsoles Martín-Santamaría,² Paul R. Crocker,⁵ Roberta Marchetti,^{1,6,*} and Alba Silipo^{1,*}

SUMMARY

Siglec-10 is an inhibitory I-type lectin selectively recognizing sialoglycans exposed on cell surfaces, involved in several patho-physiological processes. The key role Siglec-10 plays in the regulation of immune cell functions has made it a potential target for the development of immunotherapeutics against a broad range of diseases. However, the crystal structure of the protein has not been resolved for the time being and the atomic description of Siglec-10 interactions with complex glycans has not been previously unraveled. We present here the first insights of the molecular mechanisms regulating the interaction between Siglec-10 and naturally occurring sialoglycans. We used combined spectroscopic, computational and biophysical approaches to dissect glycans' epitope mapping and conformation upon binding in order to afford a description of the 3D complexes. Our outcomes provide a structural perspective for the rational design and development of high-affinity ligands to control the receptor functionality.

INTRODUCTION

Siglecs (sialic acid-binding immunoglobulin-like lectins) constitute a family of cell surface immunomodulatory receptors belonging to the immunoglobulin (Ig) superfamily that act as critical regulators of immune system (Crocker et al., 2007; Duan and Paulson, 2020; Macauley et al., 2014; Läubli and Varki, 2019) through the specific recognition of sialic acids, hallmarks of vertebrate self-glycans. Their ability to control the immune balance, triggering tolerogenic or immunogenic responses, explains why aberrant sialic acid-Siglec interactions are linked to a broad spectrum of pathologies including infection, autoimmunity, and cancer (Bochner and Zimmermann, 2015; Lübbers et al., 2018; Adams et al., 2018). Furthermore, feared human pathogens evolved the ability to shield their envelope glycans with sialic acid, mimicking Self Associated Molecular Patterns (Liu et al., 2017; Chang and Nizet, 2014) and exploiting the interaction with inhibitory Siglecs to escape host immune surveillance. In this context, the Siglec-sialoglycan axis is, therefore, an emerging attractive therapeutic target to prevent or affect the course of several immune (cancer, autoimmune, and infectious) diseases.

Siglecs are mainly expressed on innate and adaptive immune cells and display a broadly similar domain organization. These single-pass transmembrane proteins possess an extracellular N-terminal V-set domain, involved in the recognition of specific sialoglycans, and most of them encompass, in their cytoplasmic tails, one or more immunoreceptor tyrosine-based inhibitory motifs (ITIMs), involved in negative intracellular signaling pathways (Crocker et al., 2007; Duan and Paulson, 2020; Macauley et al., 2014).

Siglec-10 is an inhibitory receptor belonging to the CD33-related Siglecs family (Whitney et al., 2001) expressed on dendritic cells, human leukocytes, and B cells (Chen et al., 2009; Bandala-Sanchez et al., 2018, 2020). It is characterized by five extracellular domains and three tyrosine-based motifs in its cytoplasmic tail, exhibits only 40%–48% homology to other members of CD33-related Siglecs group, and furthermore can bind with high affinity to α -2,6- or α -2,3-sialoglycans as well as to GT1b ganglioside (Rapoport et al., 2003). Siglec-10 is the human homolog of murine Siglec-G, which is known to regulate the tolerance of the B1 sub-class of B cells, dampening autoimmune responses upon sialoglycans recognition, although with yet unknown molecular mechanisms (Chen et al., 2014). Moreover, the leading cause of bacterial gastroenteritis worldwide and a major causative trigger of the Guillain-Barré syndrome,

¹Dipartimento di Scienze Chimiche, Complesso Universitario Monte Sant'Angelo, Università di Napoli Federico II, Via Cintia 4, 80126 Napoli, Italy

²Department of Structural & Chemical Biology, Centro de Investigaciones Biológicas "Margarita Salas", CIB-CSIC, C/ Ramiro de Maeztu 9, 28040 Madrid, Spain

³Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Molecular Cell Biology and Immunology, Amsterdam Infection and Immunity Institute, Amsterdam, the Netherlands

⁴Department of Chemistry, Graduate School of Science, Osaka University, Osaka, Japan

⁵Division of Cell Signalling and Immunology, School of Life Sciences, University of Dundee, Dundee, UK

⁶Lead Contact

*Correspondence: roberta.marchetti@unina.it (R.M.), silipo@unina.it (A.S.)
<https://doi.org/10.1016/j.isci.2020.101231>



Campylobacter jejuni (Phongsisay, 2016), promotes an anti-inflammatory response mediated by its recognition by Siglec10 through the pseudaminic acid (sugar chemically related to sialic acid) decorating the bacterial flagellum (Stephenson et al., 2014). On the other hand, the interaction of Siglec-10 with the heavily glycosylated, mucin-like cell surface protein CD24, exposed on a variety of tumor cells, is an innate immune checkpoint essential for mediating anti-tumor immunity (Barkal et al., 2019; Zhang et al., 2015). Interestingly, expression of CD24 and Siglec-10 has also been observed at the fetal-maternal interface during the first 3 months of pregnancy, suggesting a possible role of Siglec-10 in establishing the immune tolerance in the placenta between the fetus and the mother (Sammar et al., 2017).

To decipher the complex mechanisms regulating host immune suppression following sialoglycan recognition by Siglec-10, the detailed knowledge of the molecular recognition and binding events represents a mandatory step. In turn, these may be translated into highly specific novel diagnostics and therapeutics, able to exploit Siglec-10 anti-inflammatory pathways for the modulation of immune cellular responses. In this context, and given the lack of any structural knowledge on Siglec-10, we used here a multidisciplinary strategy, combining spectroscopic, computational, and biophysical approaches, providing the first insights into the molecular mechanisms regulating Siglec-10 recognition of cognate (naturally occurring) sialoglycans, demonstrating the critical role played by the glycans' conformation.

RESULTS

Here, we tackled the important question of how Siglec-10 recognizes and binds, at a molecular level, different sialylated glycans, including 3'SLN and 6'SLN, **1** and **2** respectively, as well as naturally occurring complex-type *N*-glycans (Figure S1). By coupling spectroscopic approaches (such as NMR, fluorescence spectroscopy) to biochemical and computational methodologies (such as homology modeling, docking, molecular dynamics, and CORCEMA-ST calculations) we have obtained and validated 3D models of ligand-protein complexes. We therefore provide the dynamic characterization of the molecular interactions occurring in solution between Siglec-10 and cognate sialoglycans, mapping their binding epitopes and defining their bioactive conformation.

Siglec-10 Binding Specificities

The binding specificities of Siglec-10 were primarily evaluated by a solid-phase enzyme immunoassay, ELISA, and fluorescence analysis, using a recombinant soluble form of the protein in which the full extracellular region was fused to the Fc region of human IgG1. As depicted in Figure 1A, the ability of Siglec-10 to recognize terminal α 2,6- and α 2,3-linked sialic acid on various sialoglycans was demonstrated in ELISA experiments, indicating that Siglec-10 similarly recognized all the sialylated conjugates tested with a significant optical density (OD) compared with the negative control (PBS). Analogously, the fluorescence titration of increasing amounts of sialoglycans into a fixed concentration of Siglec-10 indicated that, regardless of the nature of the Neu5Ac-Gal linkage, both α 2,3- or α 2,6-linked sialoglycans could interact with the protein with binding constants in the micromolar range (Figure 1B).

Molecular Basis of Sialoglycans Recognition by Siglec-10

A combination of ligand-based NMR techniques and computational studies was employed to evaluate the epitope mapping (*group epitope mapping*) and the bioactive conformation of different sialoglycans (Figure S1) upon binding to Siglec-10.

Group Epitope Mapping and Bioactive Conformation of **1**

The saturation transfer difference (STD) NMR analysis (Marchetti et al., 2016; Meyer and Peters, 2003; Mayer and Meyer, 1999) (Figure 2) conducted on 3'-Sialyl-N-acetylglucosamine **1** in the presence of Siglec-10 confirmed the binding and allowed a detailed description of the structural elements of **1** when accommodated into the protein binding pocket (Figure 2A). The comparison of the STD NMR spectrum with the corresponding reference (off-resonance) showed that only some protons of ligand **1** exhibited STD enhancements. STD effects were calculated by use of $(I_0 - I_{\text{sat}})/I_0$, where $(I_0 - I_{\text{sat}})$ is the intensity of the relative signal in the STD NMR spectrum and I_0 is the peak intensity of an unsaturated reference spectrum (off-resonance). Siglec-10 gives the maximum saturation transfer to the acetyl group protons belonging to **K** residue of trisaccharide **1**, so this STD signal was set at 100% and the other ligand **1** protons percentages were derived accordingly. The sialic acid residue (**K**) resulted essential for the interaction with Siglec-10, giving the most significant contribution to the molecular recognition process. Indeed, the *N*-acetyl group of **K**

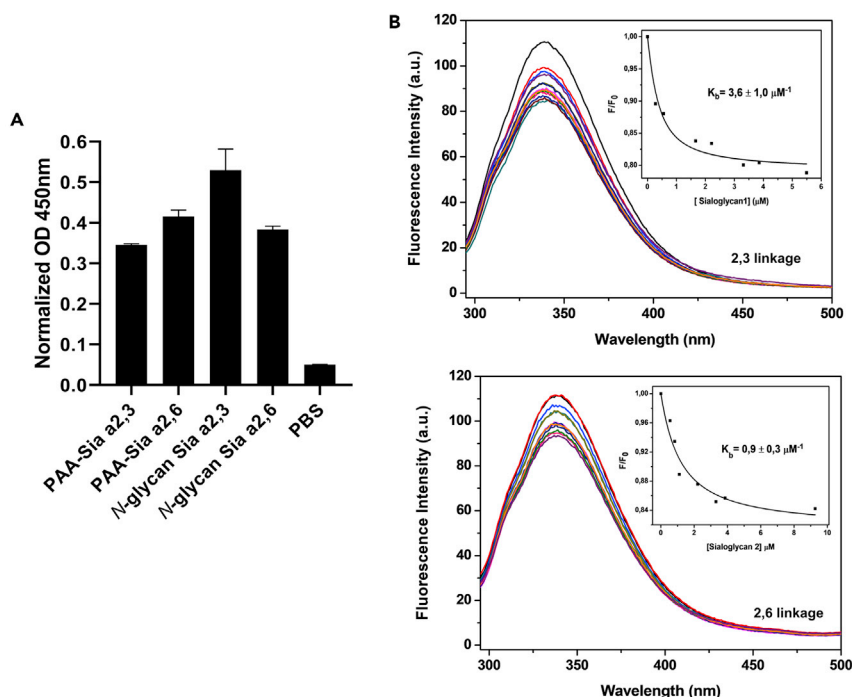


Figure 1. Siglec-10 Binding to Sialoglycans

(A) *Siglec-10 binding to sialylated conjugates detected by ELISA.* ELISA plates were coated with synthetic conjugates carrying terminal sialic acids with $\alpha 2,3$ or $\alpha 2,6$ linkages, followed by the addition of a human Fc chimera of Siglec-10. The detection of the interaction was evaluated with an anti-Human IgG-HRP (horseradish peroxidase) conjugate followed by substrate addition. The experiment has been performed three times in duplicate, showing similar results. Error bars indicate standard deviations of a representative experiment performed in duplicate. Data are normalized on BSA-treated wells used as blocking agent.

(B) *Fluorescence titration.* Fluorescence spectra of Siglec-10 (black lines) and in the presence (colored lines) of increasing amounts of the ligand sialoglycan 1 (upper panel) and 2 (lower panel), respectively. The insets report the binding isotherm and the values of the binding constants (K_D) obtained as described in the text.

showed the highest STD signal, followed by the protons signals H-6 and H-7 affected by a rather high saturation transfer, close to 80%. Significant STD enhancements, around 60%, were also observed for protons H-5, H-8 and H-9, indicating their strong involvement in the interaction, whereas weaker STD signals were observed for the diastereotopic H-3 protons of **K** and for some protons of the galactose unit **B**. Conversely, the *N*-acetylglucosamine residue (**A**) almost did not contribute to the binding process, as confirmed by the absence of any corresponding STD signal. We have also constructed the STD build-up curves (Angulo and Nieto, 2011; Yan et al., 2003) (Figure 2C, Table S1) to overcome possible artifacts due to different protons T1 relaxation times and thus to precisely map the interacting epitope of ligand 1. The strong involvement of **K** in the binding was confirmed, as well as the orientation of **B** pointing toward the protein binding site.

The determination of the epitope map of trisaccharide 1 (*group epitope mapping*) was complemented with the analysis of the conformational behavior of 1 in its free state and upon binding with Siglec-10. The accessible conformational space of 1 was accomplished by molecular mechanics (MM) and molecular dynamics (MD) simulations, together with nuclear overhauser effect (NOE)-derived experimental data, to define the preferred torsional angles ligand 1 adopted in the free and bound states. In the free state, a conformational equilibrium was observed between three main conformers, namely, *g*-/*g*+/*t*, differing by the value of the Φ (C_1 - C_2 -O- C_3') torsion angle between Sia and Gal residues ($-60^\circ/60^\circ/180^\circ$ data not shown) (Poppe et al., 1997; Forgione et al., 2020). It is worth noting that, in terms of topology, all these conformers adopt a rather extended conformation in which the sugar units fill a cone-like region of space (Xu et al., 2009) governed by a value of θ higher than 110° , where θ represents the angle between the carbon C-2 of Neu5Ac residue and C-1 atoms of the subsequent Gal and GlcNAc units (Chandrasekaran et al., 2008). During the MD simulation, ligand 1 mainly explored the *g*- ($\Phi = -60^\circ$) and *t* ($\Phi = 180^\circ$) conformations; conversely, the Ψ torsion angle (C_2 -O- C_3' -H $_3'$) remained stable along the dynamics around an average value of -11° . MD simulation results were supported by nuclear

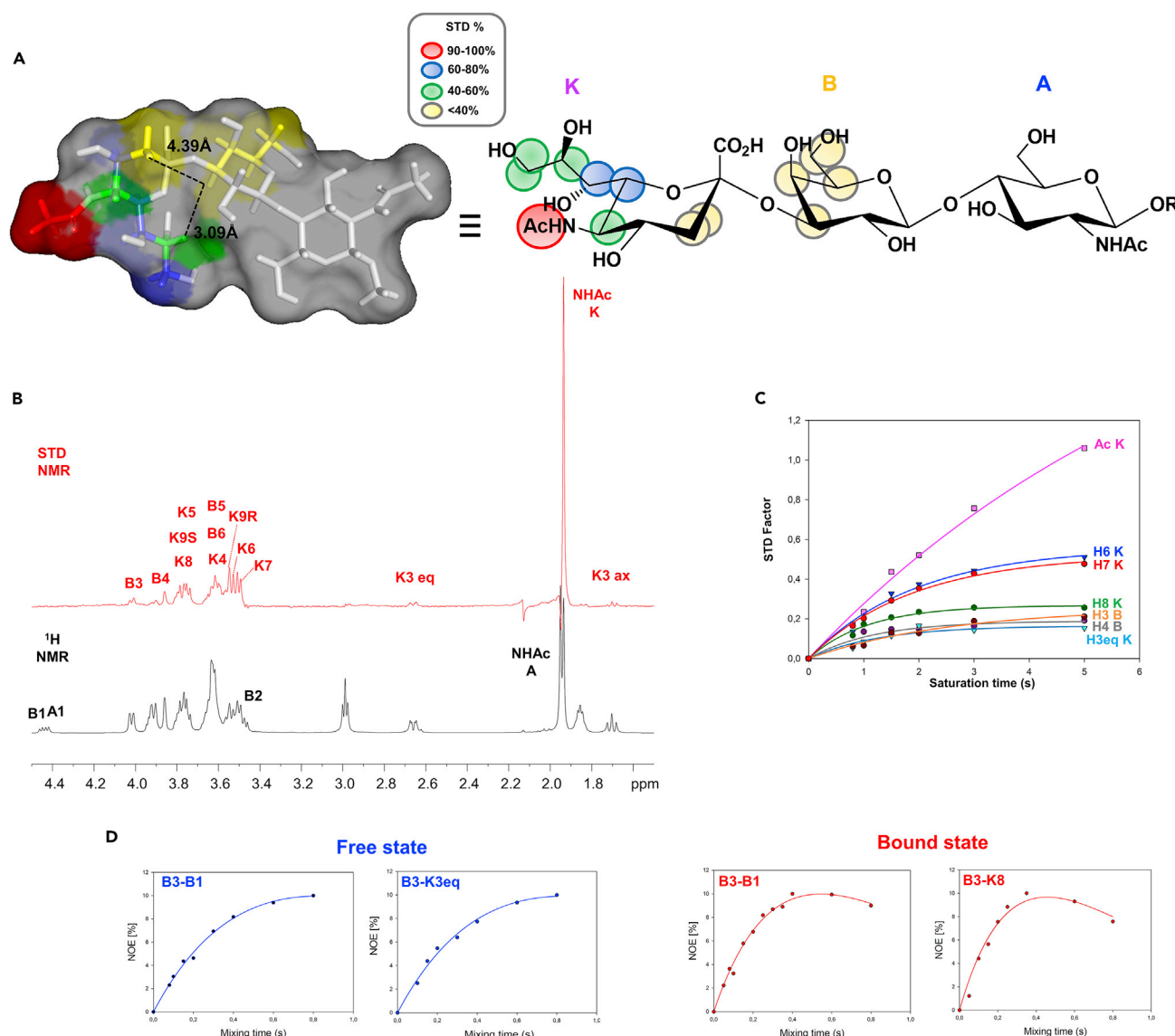


Figure 2. Analysis of the Group Epitope Mapping and Bioactive Conformation of Trisaccharide 1 upon Binding with Siglec-10

(A) STD-derived epitope mapping on the molecular envelope of ligand 1 in its bioactive minus gauche conformation.

(B) Reference ¹H NMR spectrum (black) and 1D STD NMR spectrum (red) of the Siglec-10/ligand 1 mixture (molecular ratio 1:100, T = 298 K).

(C) STD build-up curves were derived by the monoexponential equation $STD = STD_{max}[1 - \exp(-k_{sat}t)]$, constructed by fitting the STD factors at different saturation times, from 1 to 5 s. In the exponential curve, STD represents the observed signal intensity, STD_{max} is the asymptotic maximum of the curve, k_{sat} is the observed saturation rate constant, and t is the saturation time.

(D) Selected NOE build-up curves of ligand 1 in the free and bound states by selective excitation of H-3 of Gal residue.

overhauser effect spectroscopy (NOESY) NMR analysis (Marchetti et al., 2016; Poveda and Jiménez-Barbero, 1998), confirming the existence of an equilibrium between different conformations of 1 in its free state (Table 1).

As for the bound state, transferred-NOESY (tr-NOESY) NMR experiments were acquired to highlight changes in the conformation of 1 when interacting with Siglec-10. The analysis of tr-NOESY demonstrated a selection of the *g*-conformer (Φ/Ψ torsion angles of $-60^\circ/-11^\circ$) of 1 upon binding. In particular, in the tr-NOESY spectrum we observed a decrease of the cross-peak intensity of B3-K3_{ax}, an increase of B3-K8, and the absence of B3-K3_{eq} cross peak (Figure S2, left panel). The crucial *inter-residual* ¹H-¹H distances (Table 1) were extracted by construction of NOE build-up curves, clearly revealing a preference for the *g*-conformation of 1 when bound to Siglec-10 (Table 1, Figures 2C and 2D).

Distance ¹ H Neu5Ac– ¹ H Gal	Exp. Free State	Exp. Bound State	Conformer <i>g</i> - $\Phi = -60^\circ$ $\Psi = -11^\circ$	Conformer <i>g</i> + $\Phi = 60^\circ$ $\Psi = -11^\circ$	Conformer <i>t</i> $\Phi = 180^\circ$ $\Psi = -11^\circ$
H3ax Neu5Ac–H3 Gal	2.86	3.90	4.10	3.11	2.18
H3eq Neu5Ac–H3 Gal	3.65	–	4.40	2.03	3.40
H8 Neu5Ac–H3 Gal	3.41	3.19	3.36	6.17	4.31

Table 1. Conformational Behavior of 1—Key Intra-Molecular Neu5Ac-Gal Distances

The proton-proton distances were obtained by the sel-NOE experiments, acquired at different mixing times irradiating the B3 proton, using B1-B3 as reference.

Group Epitope Mapping and Bioactive Conformation of 2

The binding profile of 6'-sialyl-N-acetylglucosamine **2** (Figure S1) by Siglec-10 was investigated by means of STD NMR (Figure 3); changes in the relative enhancements and differences in the multiplicity of STD signals were detected. The epitope mapping of ligand **2** highlighted the protons specifically involved in the interaction with Siglec-10. The strongest STD signal belonged to the N-acetyl group of non-reducing Neu5Ac, K residue, providing an STD_{epitope fit} value of 100% (Table S2). As for the other proton signals belonging to sialic acid, only H-6 and H-7 showed high STD effects, thus pointing significantly close to the receptor-binding pocket, whereas the other protons exhibited STD values around or below 30%. Furthermore, the galactose moiety (B) of **2** received a significantly weaker magnetization transfer from the protein, corresponding to low STD values, and thus seemed less involved in the interaction with Siglec-10 if compared with **1**. The N-acetylglucosamine (A) of **2** revealed no STD enhancements, suggesting that it was likely solvent exposed.

The conformational behavior of **2** in its free and bound states was determined by combining NMR and computational techniques. Here, the Neu5Ac- α -(2-6)-Gal glycosidic linkage was defined by an extra torsion angle (Poppe et al., 1992), namely, ω ($O'_6-C'_6-C'_5-O'_5$), that influences the entire three-dimensional structure of **2**. In detail, three different rotamers differing in the ω value can coexist in solution, namely, *gg/tg/gt* (Patel et al., 2014) (ω $-60^\circ/180^\circ/60^\circ$, respectively), the *gt* conformer being the most populated in the free state (Di Carluccio et al., 2019). Comparing the experimental NOE-derived and theoretical distances (Table 2, Figure S2 right panel), a preference was evident for the geometry with $\phi/\psi/\omega$ torsion angles of $-60^\circ/180^\circ/60^\circ$ around Sia-Gal linkage (Family I, *gt* conformer) (Poppe et al., 1992) both in the free and bound states. The *gt* preference of **2** corresponded to a bent conformation that implicates the back-folding of the sialic acid residue and was characterized by an open umbrella-like topology (θ angle below 110° [Xu et al., 2009]) as supported by the small NOE contact between H-5 of sialic acid residue and the methyl group of N-acetylglucosamine (Figures 3C and S2 right panel).

Homology Modeling of Siglec-10

Since the crystal structure of Siglec-10 has not been solved yet, we opted to predict the protein 3D structure by means of computational methods, with the aim to describe a 3D model of Siglec-10/sialoglycans complexes. The homology model of the CRD (carbohydrate recognition domain) of Siglec-10 was derived by using Siglec-8 (PDB-ID: 2N7A), Siglec-7 (PDB-ID: 2HRL), Siglec-5 (PDB-ID: 2ZG2), and Siglec-3 (PDB-ID: 5IHB) as structural templates (see Supplemental Information for further details). The models were computed with different homology modeling servers and then refined by MD simulations as implemented within the Amber 14 package (Case et al., 2014) (Figures S3 and S4). The structural models of Siglec-10 displayed high similarity along with the corresponding templates. By comparing such models, the characteristic fold of the V-set domain of Siglecs CRD was observed, even though some differences occurred in the conformation of the variable region of Siglecs, namely, the *inter-strand* CC' loop (Figure S4).

To select the most accurate model, a detailed analysis was performed, focusing the attention on the orientation of the binding-site residues and the flanking loops. Four promising models (Figure S4) were subsequently used for docking calculations with sialoglycans by means of Autodock 4.2 (Morris et al., 2009). Selected ligand-protein complexes were then submitted to MD simulations to monitor the stability in solution (Figures S4B and S4C). The most consistent protein/ligand complexes obtained from docking and MD were subjected to CORCEMA-ST (complete relaxation and conformational exchange matrix)

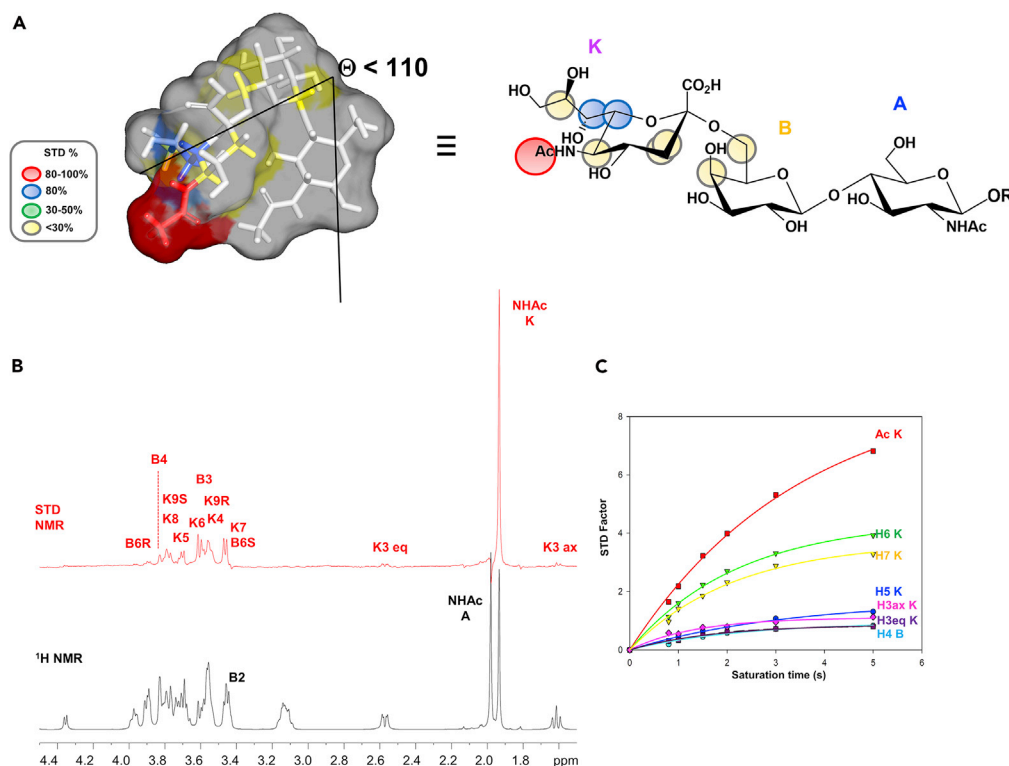


Figure 3. Analysis of the Group Epitope Mapping and Bioactive Conformation of Trisaccharide 2 upon Binding with Siglec-10

(A) Reference ^1H NMR spectrum (black) and 1D STD NMR spectrum (red) of the Siglec-10/ligand **2** mixture (molecular ratio 1:150, $T = 298\text{ K}$).

(B) STD build-up curves derived by fitting the experimental data to the monoexponential equation $\text{STD} = \text{STD}_{\text{max}}[1 - \exp(-k_{\text{sat}}t)]$.

(C) STD-derived epitope mapping on the molecular envelope of ligand **2** in its bioactive bent *gt* conformation.

calculations (Jayalakshmi and Krishna, 2002), a program that evaluates the predicted STD-NMR intensities for a proposed molecular model of a ligand-receptor complex. In order to validate the given complexes, the predicted fractional intensity changes for different ligand protons were compared with the experimental STD-NMR data through a normalized root-means square deviation (RMSD) value, also known as the NOE R factor. Thus, docking, MD simulations, and CORCEMA-ST analysis permitted the selection of *MODEL IV* as the optimal Siglec-10 model, predicted by using Siglec-7 structural template (Figure S4). It is worth to note that the selected sialoglycans/Siglec-10 complexes exhibited not only the best agreement between theoretical and experimental STD data but also the highest stability along the simulations.

3D-View of Siglec-10 in Interaction with **1**

A three-dimensional view of the Siglec-10/trisaccharide **1** complex was achieved by computational studies. Docking calculations were performed using trisaccharide **1** in each of its three possible conformations (*g*-, *g*+, *t*). Selected docked poses were then submitted to MD simulations and subjected to CORCEMA-ST calculations. The best fit between predicted and experimental data, corresponding to the lowest NOE R factor value (0.29), was given by the *MODEL IV/1* complex reported in Figure 4 (upper panel), carrying the trisaccharide **1** in the extended *g*- conformation, denoting good agreement between experimental and predicted STD data. Of note, the profiles of the theoretical STD data achieved for the other two models, bearing the trisaccharide **1** in *g* and *t* conformation, respectively, were far from the experimental STD effects, resulting in higher NOE R factor values and suggesting that these models should be excluded (Figure S6). The best selected Siglec-10/ligand **1** model (Figure 4, upper panel) displayed good stability along the MD simulations, further validating the preference of the ligand for the *g*- conformation in the bound state, as suggested by the stability of the Φ glycosidic torsion angle between the sialic acid and the galactose moiety during most of the simulation time (see Supplemental Information).

Distance	Exp Free State	Exp Bound State	Family I $\Phi = -60^\circ$ $\Psi = 180^\circ$ $\Omega = 60^\circ$	Family II $\Phi = 180^\circ$ $\Psi = 180^\circ$ $\Omega = 60^\circ$
H1 Gal -H4 GlcNAc	2.52	2.50	2.40	2.40
H3eq Neu5Ac -H3 Gal	/	/	4.93	3.84
H8 Neu5Ac -H3 Gal	/	/	4.58	3.37
H3eq Neu5Ac -H3 Gal	4.20	4.30	4.43	2.53
H8 Neu5Ac -H3 Gal	4.10	4.30	4.25	2.35
H5 Neu5Ac -CH ₃ GlcNAc	4.00	4.00	4.30	9.60

Table 2. Conformational Behavior of 2—Key Intra-Molecular Distances

Intra- and inter-residual proton-proton distances extracted from the results of the MD simulations and translated into NOE contacts according to a full-matrix relaxation approach. The corresponding average distances obtained for the simulation from $\langle r^{-6} \rangle$ values were compared with those estimated from the NOESY/tr-NOESY NMR experimental results.

Examination of the binding mode of **1** by Siglec-10 (*model IV*) revealed the typical architecture of Siglecs binding site, characterized by a solvent-exposed, highly positively charged cavity constituted by the C, F, and G β strands and surrounded by the CC' and GG' loops (Figure 4 upper panel and Figure S4). The Neu5Ac residue was placed at the top side of the G-strand, making several polar interactions with the receptor amino acids. Such interactions established by the Neu5Ac moiety were previously reported to be crucial for recognition of sialylated glycans by other CD33-related Siglecs (Crocker et al., 2007; Attil et al., 2006, 2008; Pröpster et al., 2016). During the 100ns MD simulation of the *model IV/1* complex, it was possible to observe high stability of the complex until 50 ns, without applying any restriction (Figure S5 upper panel). Alongside the simulation, ligand-protein interaction analysis identified relevant polar contacts between the trisaccharide **1** Neu5Ac unit and the residues located into the binding site of Siglec-10, namely, Arg-119, Arg-127, Asn-129, along with hydrophobic interactions involving Phe-21 and Tyr-128. In detail, the sialic acid carboxylate established a stable ionic interaction with the ammonium group from the highly conserved Arg-119. Furthermore, the hydroxyl groups of the glycerol chain of Neu5Ac formed hydrogen bonds with Asn-129 and Arg-20 residues of the protein. Also, a polar interaction between the acetamide moiety of Sia and Arg-127 residue was found. The H-bonds between the OH group at C-8 position of **1** and Asn-129 amide, and the OH group at C-9 and Arg-119 CO group were maintained along most of the simulation. The NH group of Neu5Ac acetamide moiety formed a hydrogen bond with the Arg-127 backbone, which was steadily held along the simulation. Additionally, Tyr-128 established key CH- π interactions with the C-7 and C-9 hydrogens of sialyl glycerol moiety. The only buried portion of **1** was the methyl group of Neu5Ac unit, which laid in a groove formed by Phe-21 and Tyr-128, thereby establishing CH- π interactions. In the complex, the Gal residue formed polar contacts with the binding-site residues of the protein, namely, Asn-129 and Glu-66, whereas the GlcNAc residue pointed far from the receptor surface, as confirmed by STD NMR analysis (Figure 4 upper panel and Figure S5 upper panel).

According to CORCEMA-ST analysis (Figure 4C upper panel), the protons belonging to the Neu5Ac units exhibited the highest STD effects, confirming that Neu5Ac was the major determinant of the interaction with Siglec-10. The strong STD value observed for the methyl group of the acetamide moiety of sialic acid corresponded to the close contacts between this moiety and Phe-21 and Tyr-128 side chains. Furthermore, the stable hydrogen bond established between acetamide moiety and Arg-127 backbone reinforced this interaction. Regarding STD data on the sialic acid glycerol side chain, the higher STD effect observed for the H-7 in comparison with H-8 and H-9 was ascribable to the CH- π interaction between H-7 and Tyr-128. Although the hydroxyl groups at positions 8 and 9 made direct hydrogen bonds with the backbone of Asn-129, the protons H-8 and H-9S pointed away with respect to the Tyr-128 group, consequently leading to an STD effect lower than the other Neu5Ac protons. Despite the contacts between the CC' loop of the protein with the galactose ring in trisaccharide **1** bound conformation, weaker STD effects were predicted for the Gal protons. In addition, the STD value estimated for the N-acetylglucosamine moiety were negligibly small, thus confirming that the residue was not involved in the interaction with the protein (Figure 4 upper panel).

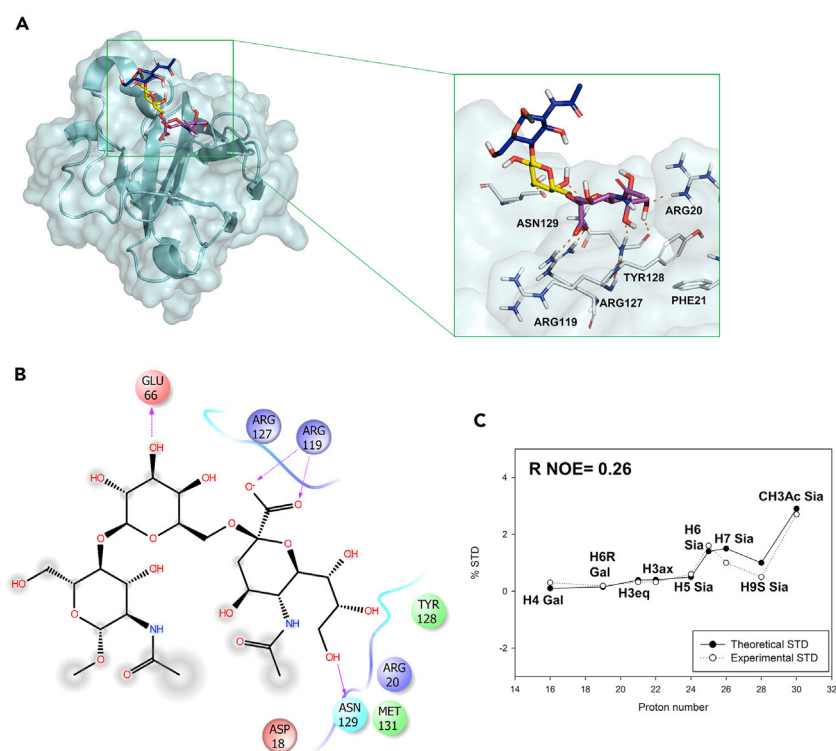
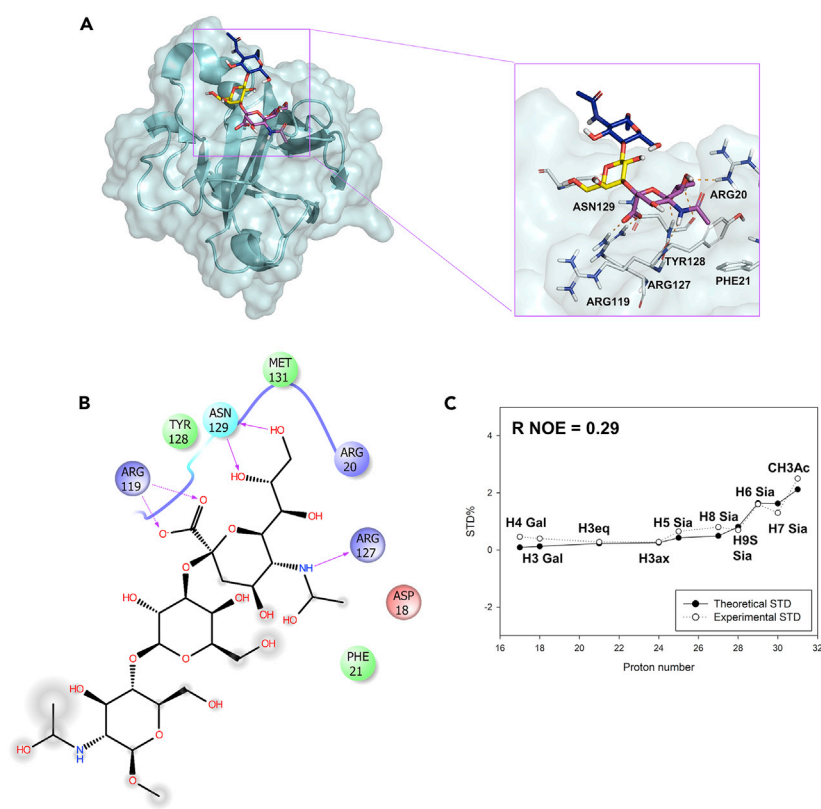


Figure 4. Docking/MD and CORCEMA-ST Analysis of 1 (Upper Panel) and 2 (Lower Panel) in the Interaction with Siglec-10

(A) Best binding pose (STD, tr-NOESY, and MD based) for the docking of the ligand bound form into binding site of Siglec-10 model derived by using Siglec-7 as structural template. The carbohydrate and interacting amino acids are shown as sticks. Dashed black lines represents hydrogen bonds.

(B) Two-dimensional plots representing the interactions between the ligand and the residues of the binding site of Siglec-10 model. Dotted arrows indicate hydrogen bonds with functional groups from side chains and solid arrows those with functional groups of the backbone. The residues shown are close to the ligand and participate in hydrophobic and polar interactions.

(C) Comparison between experimental (dashed line) and theoretical (solid line) STD data for the best model of ligand bound to the structural model of Siglec-10 CRD derived by CORCEMA-ST analysis.

3D-View of Siglec10 in Interaction with 2

To get a three-dimensional perspective of the interaction of trisaccharide **2** with Siglec-10 CRD, docking calculations were performed by using the previously selected Siglec-10 structure and following a similar protocol to that for trisaccharide **1** (see above and [Supplemental Information](#) for further details). MD simulations of the most representative complexes were performed to assess the stability of the binding poses ([Figure S4C](#)); then, with the aim of selecting the best 3D model of interaction in line with STD data, CORCEMA-ST calculations were carried out. The lowest NOE R factor, with the value of 0.26, was obtained for the complex reported in the lower panel of [Figure 4](#) in which the trisaccharide **2** adopted a *gt* conformation (see also [Figure S6B](#)).

The terminal sialic acid was again the major determinant of ligand binding, establishing several polar interactions with the protein, which were comparable with those observed for the trisaccharide *MODEL IV/1* complex ([Figure 4](#) upper panel). As in trisaccharide **1**, the methyl group of Neu5Ac lies in the hydrophobic groove defined by Phe-21 and Tyr-128 ([Figure 4](#) lower panel). Furthermore, the Neu5Ac residue of **2** formed the key salt bridge between its carboxylate and the guanidinium group of Arg-119; hydrogen bonds between the backbone of Asn-129 and the hydroxyl groups of the glycerol chain, namely, OH-8 and OH-9, were also observed. Notably, these interactions remained stable along the MD simulation. As observed for trisaccharide **1**, a direct involvement of the Gal unit of **2** in the binding was evident with the hydrogen bond occurring between the Gal hydroxyl group in position 3 and the Glu-66 residue. Furthermore, the protein did not engage interactions with the GlcNAc unit of **2**.

From the MD simulation performed on the selected *MODEL IV/2* complex, it was possible to observe that, although **2** remained steadily anchored to the protein via the sialic acid moiety, the bound ligand explored both *gt* and *tg* conformations along the dynamic ([Figure S5](#) lower panel). The ω angle, indeed, fluctuated between $\omega = 60^\circ$ and $\omega = 180^\circ$. This suggested that ligand **2** displayed higher conformational flexibility, even upon binding, with respect to ligand **1**. However, from the CORCEMA-ST analysis, the best fit between experimental and theoretical data was obtained for the *gt* conformer ([Figure 4C](#) lower panel). In detail, the highest predicted STD value was given by the methyl group of the *N*-acetamide moiety of Neu5Ac. This could be ascribable to the vicinity of the acetamide group toward the aromatic residues Phe-21 and Tyr-129. For the same reasons discussed for **1**, the H-7 of the sialic acid glycerol side chain of trisaccharide **2** displayed higher STD effects with respect to the protons in position 8 and 9. Despite the proximity of the CC' loop of the protein to the galactoside ring of the trisaccharide **2**, lower STD effects were observed for the galactose moiety of **2**. In complete agreement with the experimental STD data, no significant STD effects were predicted for the GlcNAc moiety ([Figure 4](#) lower panel).

Molecular Interaction between Siglec-10 and Complex-type N-glycans

The interaction between naturally occurring glycans and Siglec-10 has also been investigated in order to unveil if and how a different presentation of the glycan epitope could influence its recognition by the receptor protein. Two biantennary complex-type undecasaccharides-Asn (ligands **3** and **4**, [Figure S1](#)), bearing α 2,3- and α 2,6-linked sialic acid at glycan antennae termini, were used. Interestingly, NMR and computational methods suggested that only the outermost portion of the complex glycoconjugates was accommodated in the binding pocket of the protein with the inner residues not participating in the molecular recognition ([Figure 5](#)). Analysis of STD NMR spectra acquired in the presence of Siglec-10 ([Figure S7](#)) clearly revealed the key role of the sialic acid in the recognition and binding process, and the epitope mapping of glycans **3** and **4** was indeed comparable with those obtained with the smaller ligands **1** and **2**; the acetyl group of the sialic acid displayed the strongest STD intensity and other STD enhancements

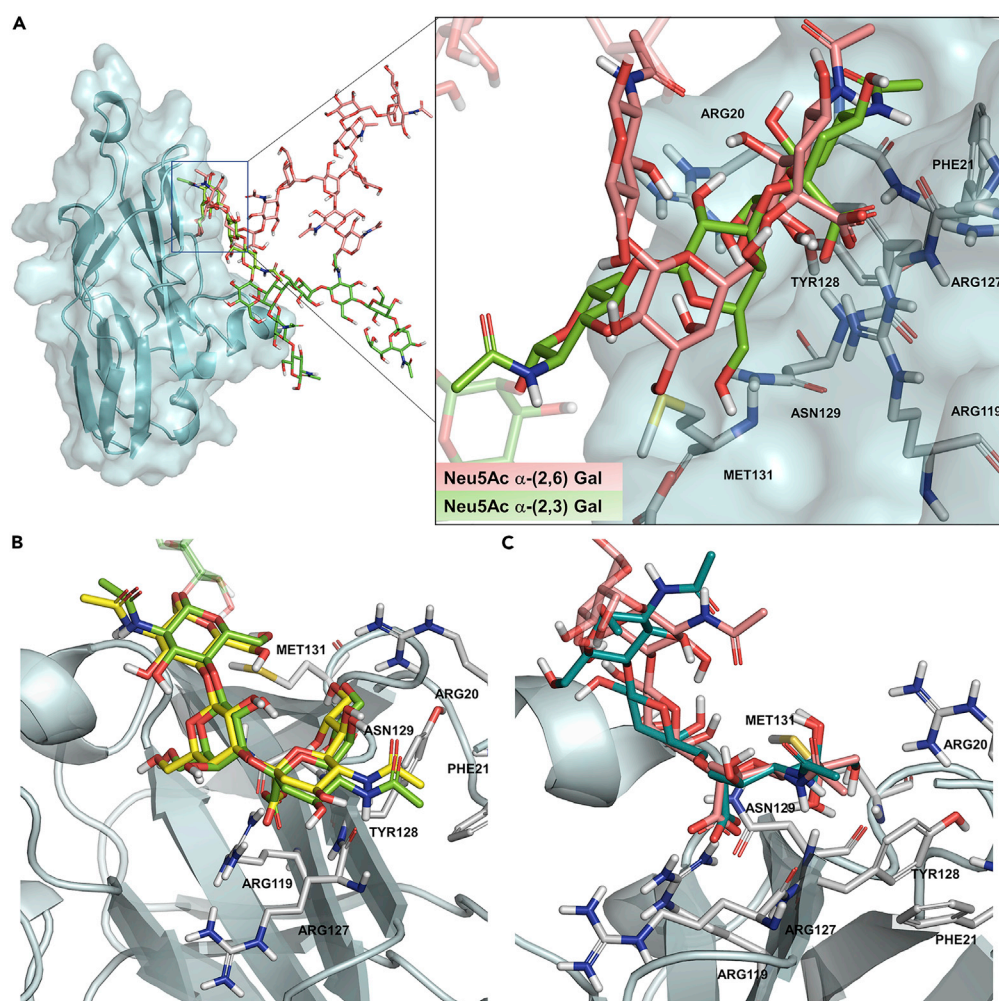


Figure 5. Molecular Recognition of Longer N-glycans by Siglec-10

(A) Full view of the 3-D complexes of ligand 3 (green) and ligand 4 (pink) using Siglec-10 (*model IV*) structural template obtained by docking calculations.

(B) Superimposition of ligand 1 (yellow) and ligand 3 (green) in the Siglec-10 (*model IV*) binding site.

(C) Superimposition of ligand 2 (blue) and ligand 4 (pink) in the Siglec-10 (*model IV*) binding site.

corresponded to H-5/H-9 protons of the Neu5Ac moiety. Furthermore, STD effects for H-4 and H-6 of the galactose unit, in the case of the α 2,6 undecasaccharide, and for H-3/H-6, in the case of the α 2,3-linked sialoglycan, were observed. In addition, the protons belonging to the other sugar residues of the branches and the N-glycan pentasaccharide core did not appear in the STD spectra, clearly suggesting that the recognition was mainly driven by the Neu5Ac-Gal moieties. The other sugar residues were indeed solvent exposed and far from the protein surface.

These results were further confirmed by modeling 3 and 4 into the Siglec-10 binding site (Figure 5). As shown in the resulting three-dimensional structures of the corresponding complexes, good predicted binding poses were obtained for both ligands, confirming that longer complex-type glycans can be accommodated in the protein binding pocket, displaying a similar network of interactions with respect to those established by sialoglycans 1 and 2 upon protein binding.

In detail, as expected, both complex glycoconjugates interacted with Siglec-10 binding-site residues mostly via the Neu5Ac moiety. Indeed, the ionized carboxylic group of sialic acid formed a polar interaction with the lateral chain of Arg-119 and Tyr-128 contributed significantly to the recognition establishing CH- π contacts with the Neu5Ac glycerol chain.

In the case of sialoglycan **3**, the hydroxyl groups of Neu5Ac lateral chain were also involved in hydrogen bonds with Arg-127 and Arg-20. The Sia acetamide moiety formed a hydrogen bond with the Arg-127 backbone, whereas the Phe-21 residue on Siglec-10 was responsible for the hydrophobic interaction with Neu5Ac acetyl group. The Gal moiety was also placed inside the binding pocket, making polar contacts with Glu-66 and Asn-129 residues. A hydrogen bond between Asn-129 and the hydroxyl group at C-9 of Neu5Ac was instead observed for sialoglycan **4** in the interaction with Siglec-10. As in **3**, the Gal moiety of **4** was in close proximity to the receptor surface, making polar contacts with Glu-66 residue. In agreement with NMR results, in both 3D complexes, the inner residues of the oligosaccharides were far from the receptor surface and did not directly contribute to the binding. Of note, the conformation of the sialylated lactosamine branches of **3** and **4**, involved in the interaction with Siglec-10, was comparable with that adopted by the sialoglycans **1** and **2**, respectively, upon binding. Moreover, independently from the sialic acid linkage to the adjacent galactose moiety, both complex glycans preferentially exhibited an extended conformation when placed in the binding site of the protein, although a certain flexibility around the Man- α (1,6)-Man glycosidic linkage was observed (Figure 5).

Conclusions

In summary, the interaction of Siglec-10 with differently sialylated glycans was here investigated by an integrated approach based on NMR, docking, and molecular modeling. Although the crystal structure of Siglec-10 has not been solved yet, the orthogonal methods applied here allowed a molecular description of the recognition and binding processes driving Siglec-10-sialoglycans complexes, providing unique information about dynamics and molecular interactions. The ligands' epitopes in their bound conformations were indeed profiled and consistent 3D models of the interactions were presented.

A comparison of the epitope maps obtained for α 2,3- and α 2,6-sialoglycans, which constitute the terminal end of the canonical sialylated complex-type *N*-glycans exposed on mammalian cells, highlighted that in each case the binding process to Siglec-10 was mainly driven by the sialic acid moiety, giving a 100% STD_{epitope fit}, and galactose unit, whereas the *N*-acetylglucosamine pointed away from Siglec-10 binding pocket. On the other hand, some differences were evident from the STD-derived epitope maps, indicating that ligand **2** established less contacts within the protein binding pocket, as demonstrated by the lower involvement in the interaction of protons H-5, H-8, and H-9 of sialic acid and of the adjacent galactose unit.

The bound conformations of α 2,3- and α 2,6-sialoglycans were obtained by combining NOE-derived data with MM and MD simulations. The *g*- conformer of the α 2,3-sialoglycans was selected upon binding to Siglec-10; instead, the α 2,6-sialoglycans mainly assumed a bent *gt* conformation both in the free and bound states. Thus, our data suggested that, upon binding, the terminal portion of sialylated glycan chains preferentially adopted a cone-like or an umbrella-like topology depending on the sialic acid linkage to the underlying galactose residue.

Extensive homology modeling studies provided a 3-D model of human Siglec-10. Then, further computational approaches, namely docking and MD studies, together with CORCEMA-ST protocol, were used to either indirectly support the reliability of the 3-D model of Siglec-10 and to obtain structural insights into the molecular basis of *N*-glycan recognition by Siglec-10. Our three-dimensional models of the Siglec-10/sialoglycan complexes highlighted the architecture of the binding site essentially composed of hydrophobic (Phe-21), polar (Tyr-128, Asn-129), basic (Arg-119, Arg-127), and acid (Glu-66) residues. In detail, the Arg-119, Arg-128, Asn-129 residues were directly involved in the interaction with sialic acid, suggesting they shaped a region relevant for the molecular recognition, matching those previously reported for other studied members of the Siglec family (Crocker et al., 2007). Moreover, we observed that the CC' loop belonging to the Siglec-10 V-set domain established key interactions with the adjacent galactose moiety that appears to participate in the interaction process, whereas the GlcNAc residue showed no relevant contribution to the binding.

The ability of Siglec-10 to recognize both α -2,3 and α -2,6 linkages to a similar extent indicates a lower specificity (Macauley et al., 2014; Blixt et al., 2003) compared with other homologous Siglecs, likely due to a wider binding site. The CC' loop region predicted for Siglec-10, indeed, points outward to the binding residues, in contrast to some other CD33-like Siglecs (for example, Siglec-5, Siglec-3, and Siglec-8) but as already reported for Sialoadhesin, Siglec-4, and Siglec-7 (Figure 6). This could account for the greater capacity of the protein to flexibly accommodate sialoglycans with different shapes and conformation, α 2,3-sialoglycan **1** in

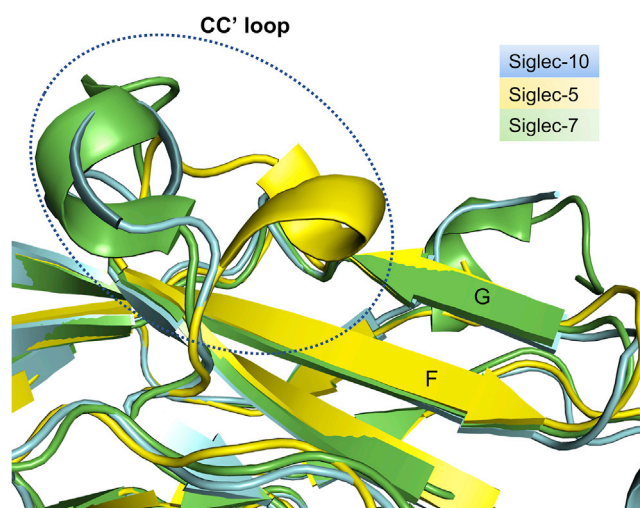


Figure 6. Siglec-10 CC' Loop Shape

Comparison of Siglec-10 (*model IV*) loop region with other CD33-like Siglec, namely, Siglec-5 (PDB ID: 2zg2) and Siglec-7 (PDB ID: 2hrl). According to the predicted model, Siglec-10 exhibits an “open” loop, similarly to Siglec-7.

its extended topology and the α 2,6-sialoglycan **2** in a more compact bent conformation. Furthermore, also longer biantennary glycoconjugates, ligands **3** and **4** (Figure S1), could be accommodated in the protein binding pocket where the terminal ends preferentially adopted a cone-like or an umbrella-like topology respectively. Their outermost region was directly involved in polar and hydrophobic interactions with amino-acidic residues, whereas the inner moieties of the oligosaccharides pointed away from the protein surface. Of note, for both α 2,3- and α 2,6-linked complex type *N*-glycans **3** and **4**, an extended conformation of the entire undecasaccharide seemed to be preferred (Figure 5) and, interestingly, the topology of the terminal sialylated portions oriented differently the long saccharide chains with respect to the binding site.

Given the pivotal roles that Siglecs play in the regulation of immune cell functions, they have been widely recognized as targets for the development of immunotherapeutics against a broad range of diseases. The receptor-sialoglycan interactions can mediate events like cell adhesion, signaling and interaction, inhibition or regulation of the immune cell activation, and uptake of sialylated pathogens. The disturbance of the sialic acid-Siglec axis can contribute to the onset and development of diseases such as autoimmunity, infection, inflammation, aging, and cancer. Worthy of note is the role played by the inhibitory receptor Siglec-10 in several crucial patho-physiological processes. Siglec-10 is engaged by sialylated glycoproteins, as tumor-expressed CD24, to promote the tumor immune evasion, or by dangerous pathogens, including *Campylobacter*, to subvert the host immune system. Although our understanding of the biological roles of Siglecs has been expanding over the past decade, a detailed molecular knowledge of their interaction with glycan ligands is still an unmet challenge. For instance, despite Siglec-10 is an appealing target for the development of glyco-based immunotherapy (Angata et al., 2015), neither the three-dimensional structure of protein nor the molecular basis of its interactions with complex glycans had been previously unraveled. Thus, the outcomes shown here have improved our knowledge of the molecular mechanisms occurring between Siglec-10 and sialylated glycans, providing a structural point of view for the design and development of high-affinity ligands, in their bioactive conformations, able to control the receptor functionality.

Limitation of the Study

This study describes the interactions of the biologically relevant protein Siglec-10 with naturally occurring sialoglycans by using different biophysical techniques, including ELISA, steady-state fluorescence, and mainly NMR spectroscopy, combined with computational chemistry and CORCEMA-ST calculations. In the absence of high-resolution structural information on the receptor Siglec-10, we resorted to homology modeling following a convergent approach of combining different templates for the homology calculation and different modeling approaches to validate the generated models (docking, MD simulations, CORCEMA-ST). However, the structural conclusions we derived from homology modeling will require the definition of the three-dimensional structure of the protein Siglec-10 to avoid misinterpretation of the data.

Further development of this work may also include the search of natural ligands for Siglec-10. The functional significance of the Siglec-10/CD24 interaction has been previously shown; however, other heavily *N*-glycosylated and/or mucin-type glycoproteins could be further candidate ligands involved in the interaction with Siglec-10.

Resource Availability

Lead Contact

Roberta Marchetti (roberta.marchetti@unina.it).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

This study did not generate datasets.

METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101231>.

ACKNOWLEDGMENTS

The authors sincerely thank GlyTech, Inc. for providing *N*-glycan ligand **3** and **4**. This study was supported by the project “GLYTUNES” funded by MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN 2017) (2017XZ22BK, 2019–2022) to A.S.; by progetto POR SATIN and Progetto POR Campania Oncoterapia to A.M.; by the European Commission (H2020-MSCA- 814102-SWEET CROSSTALK project) to A.M., R.M., and A.S.. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement No 851356 to R.M. FSE, PON Ricerca e Innovazione 2014–2020, Azione I.1 “Dottorati Innovativi con caratterizzazione Industriale” is acknowledged for funding the PhD grant to R.E.F. Grants by the Spanish Ministry of Science MICINN (CTQ2017-88353-R and fellowship BES 2015–071588 to J.G.-C.) and Wellcome Trust 103744/Z/14/Z to P.R.C. are acknowledged.

AUTHOR CONTRIBUTIONS

A.S. and R.M. conceived and designed the project. R.M., R.E.F., C.D.C., A.S., and A.M. carried out NMR, docking, and CORCEMA experiments and analyzed the results. R.E.F., J.G.-C., and S.M.-S. performed homology modeling and MD simulations. P.R.C., A.A., and R.G. produced the protein. F.C. performed ELISA experiments. P.D.V. and F.B. carried out fluorescence analysis. K.F. and Y.M. provided the glycans. All the authors wrote, revised, and reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 17, 2020

Revised: May 20, 2020

Accepted: May 29, 2020

Published: June 26, 2020

REFERENCES

Adams, O.J., Stanczak, M.A., von Gunten, S., and Läubli, H. (2018). Targeting sialic acid-Siglec interactions to reverse immune suppression in cancer. *Glycobiology* 28, 640–647.

Angata, T., Nychola, C.M., and Macauley, M.S. (2015). Therapeutic targeting of siglecs using antibody- and glycan-based approaches. *Trends Pharmacol. Sci.* 36, 645–660.

Angulo, J., and Nieto, P.M. (2011). STD-NMR: application to transient interactions between biomolecules—a quantitative approach. *Eur. Biophys. J.* 40, 1357–1369.

- Attrill, H., Takazawa, H., Witt, S., Kelm, S., Isecke, R., Brossmer, R., Ando, T., Ishida, H., Kiso, M., Crocker, P.R., and van Aalten, D.M. (2006). The structure of siglec-7 in complex with sialosides: leads for rational structure-based inhibitor design. *Biochem. J.* 397, 271–278.
- Attrill, M.A., Trandem, K., and Sun, P.D. (2008). Structural implications of siglec-5-mediated sialoglycan recognition. *J. Mol. Biol.* 375, 437–447.
- Bandala-Sanchez, E., Bediaga, G.N., Goddard-Borger, E.D., Ngui, K., Naselli, G., Stone, N.L., Neale, A.M., Pearce, L.A., Wardak, A., Czabotar, P., et al. (2018). CD52 glycan binds the proinflammatory B box of HMGB1 to engage the Siglec-10 receptor and suppress human T cell function. *Proc. Natl. Acad. Sci. U S A* 115, 7783–7788.
- Bandala-Sanchez, E., Bediaga, N.G., Naselli, G., Neale, A.M., and Harrison, L.C. (2020). Siglec-10 expression is up-regulated in activated human CD4⁺ T cells. *Hum. Immunol.* <https://doi.org/10.1016/j.humimm.2020.01.009>.
- Barkal, A.A., Brewer, R.E., Markovic, M., Kowarsky, M., Barkal, S.A., Zaro, B.W., Krishnan, V., Hatakeyama, J., Dorigo, O., Barkal, L.J., et al. (2019). CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy. *Nature* 572, 392–396.
- Blixt, O., Collins, B.E., van den Nieuwenhof, I.M., Crocker, P.R., and Paulson, J.C. (2003). Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein. *J. Biol. Chem.* 278, 31007–31019.
- Bochner, B.S., and Zimmermann, N. (2015). Role of siglecs and related glycan-binding proteins in immune responses and immunoregulation. *J. Allergy Clin. Immunol.* 135, 598–608.
- Di Carluccio, C., Crisman, E., Manabe, Y., Forgione, R.E., Lacetera, A., Amato, J., Pagano, B., Randazzo, A., Zampella, A., Lanzetta, R., et al. (2019). Characterisation of the dynamic interactions between complex N-glycans and human CD22. *ChemBioChem.* <https://doi.org/10.1002/cbic.201900295>.
- Case, D.A., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, T.E., Darden, T.A., Duke, H.G., Goetz, A.W., Gusarov, S., et al. (2014). AMBER 14 (University of California, San Francisco).
- Chandrasekaran, A., Srinivasan, A., Raman, R., Viswanathan, K., Raguram, S., Tumpey, T.M., Sasisekharan, V., and Sasisekharan, R. (2008). Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat. Biotechnol.* 26, 107–113.
- Chang, Y.C., and Nizet, V. (2014). The interplay between Siglecs and sialylated pathogens. *Glycobiology* 24, 818–825.
- Chen, G.Y., Tang, J., Zheng, P., and Liu, Y. (2009). CD24 and Siglec-10 selectively repress tissue damage-induced immune responses. *Science* 323, 1722–1725.
- Chen, G.-Y., Brown, N.K., Zheng, P., and Liu, Y. (2014). Siglec-G/10 in self–nonself discrimination of innate and adaptive immunity. *Glycobiology* 24, 800–806.
- Crocker, P.R., Paulson, J.C., and Varki, A. (2007). Siglecs and their roles in the immune system. *Nat. Rev. Immunol.* 7, 255–256.
- Duan, S., and Paulson, J.C. (2020). Siglecs as immune cell checkpoints in disease. *Annu. Rev. Immunol.* <https://doi.org/10.1146/annurev-immunol-102419-035900>.
- Forgione, R.E., Di Carluccio, C., Kubota, M., Manabe, Y., Fukase, K., Molinaro, A., Hashiguchi, T., Marchetti, R., and Silipo, A. (2020). Structural basis for Glycan-receptor binding by mumps virus hemagglutinin-neuraminidase. *Sci. Rep.* 10, 1589.
- Jayalakshmi, V., and Krishna, N.R. (2002). Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer effects in reversibly forming ligand-receptor complexes. *J. Magn. Reson.* 155, 106–118.
- Läubli, H., and Varki, A. (2019). Sialic acid-binding immunoglobulin-like lectins (Siglecs) detect self-associated molecular patterns to regulate immune responses. *Cell. Mol. Life Sci.* 77, 593–605.
- Liu, Y.C., Yu, M.M., Chai, Y.F., and Shou, S.T. (2017). Sialic acids in the immune response during sepsis. *Front. Immunol.* 8, 1601.
- Lübbbers, J., Rodríguez, E., and van Kooyk, Y. (2018). Modulation of immune tolerance via siglec-sialic acid interactions. *Front. Immunol.* 9, 2807.
- Macauley, M.S., Crocker, P.R., and Paulson, J.C. (2014). Siglec-mediated regulation of immune cell function in disease. *Nat. Rev. Immunol.* 14, 653–666.
- Marchetti, R., Perez, S., Arda, A., Imbert, A., Jimenez-Barbero, J., Silipo, A., and Molinaro, A. (2016). “Rules of Engagement” of proteins-glycoconjugates interactions: a molecular view achievable by NMR spectroscopy and molecular modeling. *ChemistryOpen* 5, 274–296.
- Mayer, M., and Meyer, B. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* 38, 1784–1788.
- Meyer, B., and Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.* 42, 864–890.
- Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791.
- Patel, D.S., Pendrill, R., Mallajosyula, S.S., Widmalm, G., and MacKerell, A.D., Jr. (2014). Conformational properties of α - or β -(1→6)-linked oligosaccharides: Hamiltonian replica exchange MD simulations and NMR experiments. *J. Phys. Chem. B* 118, 2851–2871.
- Phongsisay, V. (2016). The immunobiology of *Campylobacter jejuni*: innate immunity and autoimmune diseases. *Immunobiology* 221, 535–543.
- Poppe, L., Struike-Prill, R., Meyer, B., and van Halbeek, H. (1992). The solution conformation of sialyl-et(2 6)-lactose studied by modern NMR techniques and Monte Carlo simulations. *J. Biomol. NMR* 2, 109–136.
- Poppe, L., Brown, G.S., Philo, J.S., Nikrad, P.V., and Shah, B.H. (1997). Conformation of sLex Tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. *J. Am. Chem. Soc.* 119, 1727–1736.
- Poveda, A., and Jiménez-Barbero, J. (1998). NMR studies of carbohydrate–protein interactions in solution. *Chem. Soc. Rev.* 27, 133–144.
- Pröpster, J.M., Yang, F., Rabbani, S., Ernst, B., Allain, F.H.-T., and Schubert, M. (2016). Structural basis for Siglec-8-glycan recognition. *Proc. Natl. Acad. Sci. U S A* 113, E4170–E4179.
- Rapoport, E., Mikhalyov, I., Zhang, J., Crocker, P., and Bovin, N. (2003). Ganglioside binding pattern of CD33-related siglecs. *Bioorg. Med. Chem. Lett.* 13, 675–678.
- Sammar, M., Siwetz, M., Meiri, H., Fleming, V., Altevogt, P., and Huppertz, B. (2017). Expression of CD24 and Siglec-10 in firsttrimesterplacenta: implications for immune tolerance at the fetal-maternal interface. *Histochem. Cell Biol.* 147, 565–574.
- Stephenson, H.N., Mills, D.C., Jones, H., Milioris, E., Copland, A., Dorrell, N., Wren, B.W., Crocker, P.R., Escors, D., and Bajaj-Elliott, M. (2014). Pseudaminic acid on *Campylobacter jejuni* flagella modulates dendritic cell IL-10 expression via Siglec-10 receptor: a novel flagellin–host interaction. *J. Infect. Dis.* 210, 1487–1498.
- Whitney, G., Wang, S., Chang, H., Cheng, K.Y., Lu, P., Zhou, X.D., Yang, W.P., McKinnon, M., and Longphre, M. (2001). A new siglec family member, siglec-10, is expressed in cells of the immune system and has signaling properties similar to CD33. *Eur. J. Biochem.* 268, 6083–6096.
- Xu, D., Newhouse, E.I., Amaro, R.E., Pao, H.C., Cheng, L.S., Markwick, P.R., McCammon, J.A., Li, W.W., Arzberger, P.W., and Distinct. (2009). Glycan topology for avian and human sialopentasaccharide receptor analogues upon binding different hemagglutinins: a molecular dynamics perspective. *J. Mol. Biol.* 387, 465–491.
- Yan, J., Kline, A.D., Mo, H., Shapiro, M.J., and Zartler, E.R. (2003). The effect of relaxation on the epitope mapping by saturation transfer difference NMR. *J. Magn. Reson.* 163, 270–276.
- Zhang, P., Lu, X., Tao, K., Shi, L., Li, W., Wang, G., and Wu, K. (2015). Siglec-10 is associated with survival and natural killer cell dysfunction in hepatocellular carcinoma. *J. Surg. Res.* 194, 107–113.

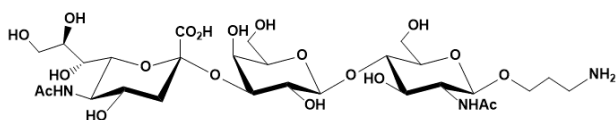
Supplemental Information

Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10

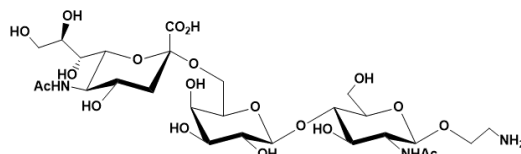
Rosa Ester Forgione, Cristina Di Carluccio, Juan Guzmán-Caldentey, Rosa Gaglione, Filomena Battista, Fabrizio Chiodo, Yoshiyuki Manabe, Angela Arciello, Pompea Del Vecchio, Koichi Fukase, Antonio Molinaro, Sonsoles Martín-Santamaría, Paul R. Crocker, Roberta Marchetti, and Alba Silipo

Supplemental Information

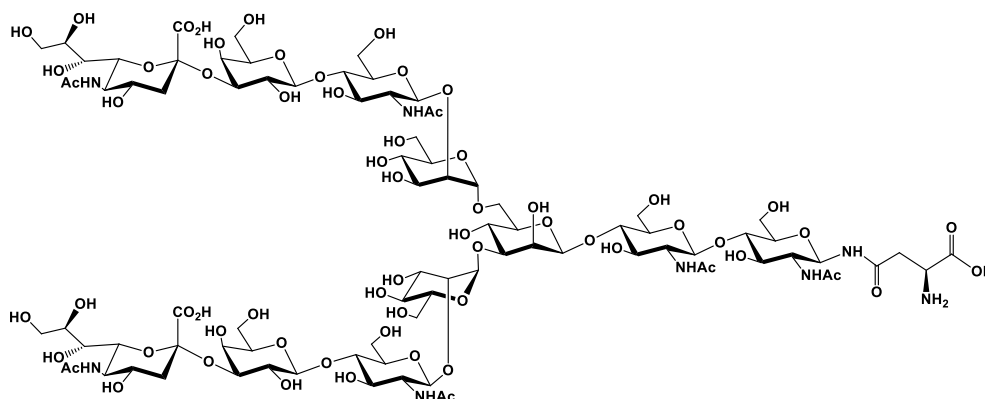
Ligand 1



Ligand 2



Ligand 3



Ligand 4

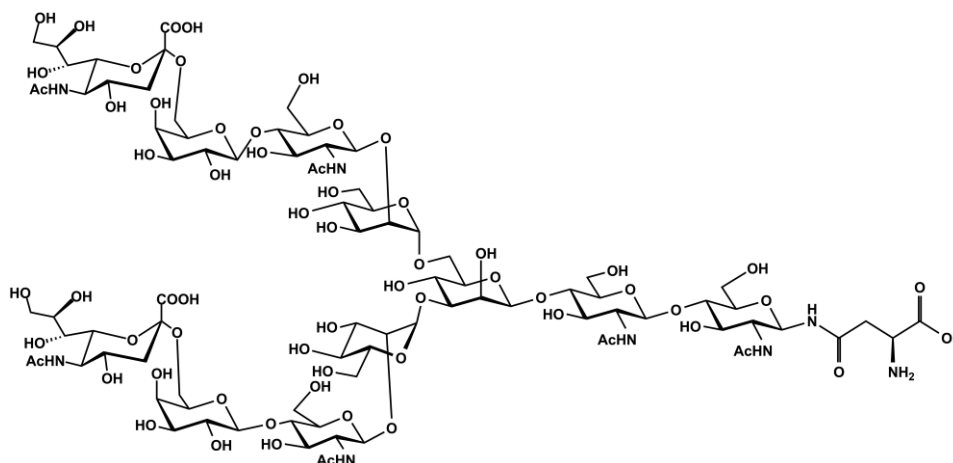


Figure S1 Related to Figures 1 and 5. Structures of sialylated ligands used in this study. The two trisaccharides, Neu5Ac- α -(2 \rightarrow 3)-Gal- β -(1 \rightarrow 4)-GlcNAc, ligand 1, and Neu5Ac- α -(2 \rightarrow 6)-Gal- β -(1 \rightarrow 4)-GlcNAc, ligand 2 differ in the type of glycosidic linkage at the non-reducing end; the longer glycans, ligands 3 and 4, are commonly exposed on mammalian cells.

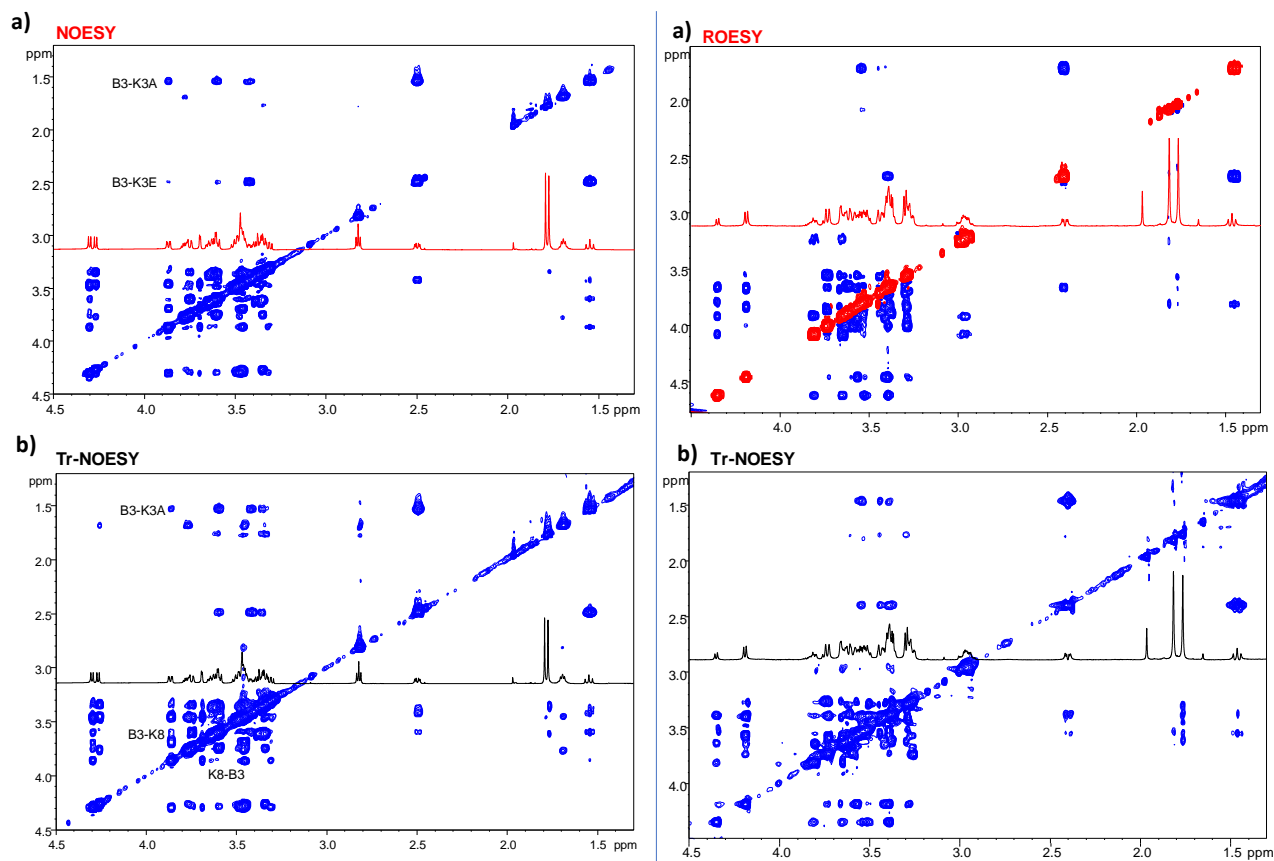


Figure S2 Related to Figures 2 and 3. Conformational analysis of ligands 1 and 2. Left panel: NOESY spectrum (a) of ligand 1 in its free state (mixing time of 600ms) and tr-NOESY spectrum (b) of ligand 1 bound to Siglec-10 (mixing time of 400ms) at 283K. The protein-ligand molar ratio was set at 1:10. Some differences in terms of signal intensities were observed comparing the free and bound state. **Right panel:** ROESY spectrum (a) of ligand 2 in its free state (mixing time of 450ms) and tr-NOESY spectrum (b) of ligand 2 bound to Siglec-10 (mixing time of 400ms) at 283K. The protein-ligand molar ratio was set at 1:10.

Table S1 Related to Figure 2. Experimental STD epitopes of trisaccharide 1 bound to Siglec-10.

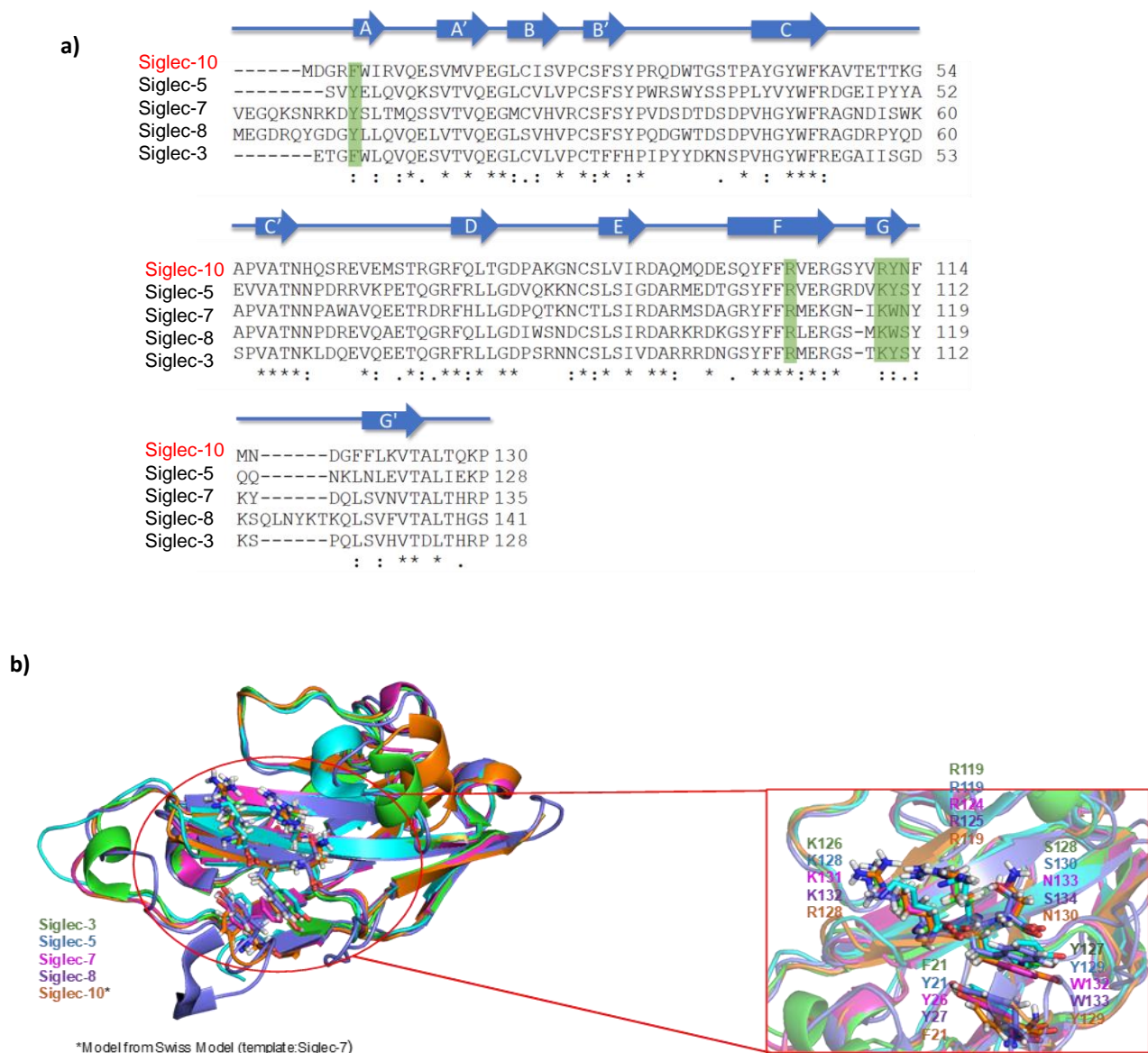
STD intensities of ligand 1 bound to Siglec-10 experimentally calculated at different saturation times, using the monoexponential equation $STD(t_{sat}) = STD_{max} \times [1 - \exp(-k_{sat} \times t_{sat})]$.

¹ H	STD _{max}	K _{sat}	STD (fit)	STD epitopes (fit)
Ac Neu5Ac	1.8331	0.1871	0.342973	100%
H6 Neu5Ac	0.5615	0.5144	0.288836	84.2%
H7 Neu5Ac	0.5251	0.5249	0.275624	80.4%
H8 Neu5Ac	0.263	0.7105	0.186862	54.4%
H9S Neu5Ac	0.3128	0.5849	0.182957	53.3%
H4 Gal	0.2017	0.5974	0.120496	35.1%
H3 Gal	0.2919	0.3848	0.112323	32.7%
H5 Neu5Ac	0.2243	0.6552	0.146961	42.8%
H3eq Neu5Ac	0.1651	0.7792	0.128646	37.5%

Table S2 Related to Figure 3. Experimental STD epitopes of trisaccharide 2 bound to Siglec-10.

STD intensities of ligand 2 bound to Siglec-10 experimentally calculated at different saturation times, using the monoexponential equation $STD(t_{sat}) = STD_{max} \times [1 - \exp(-k_{sat} \times t_{sat})]$.

¹ H	STD _{max}	K _{sat}	STD (fit)	STD epitopes (fit)
KAc Neu5Ac	9.0345	0.2863	2.586577	100%
H6 Neu5Ac	4.4377	0.4472	1.984539	76.7%
H7 Neu5Ac	3.7108	0.4639	1.72144	66.6%
H3eq Neu5Ac	0.8736	0.5546	0.484499	18.7%
H3ax Neu5Ac	1.1098	0.7396	0.820808	31.7%
H9S Neu5Ac	1.4138	0.5461	0.772076	29.8%
H5 Neu5Ac	1.684	0.3123	0.525913	20.3%
H6R Gal	0.5481	0.7821	0.428669	16.6%
H4 Gal	0.9568	0.4301	0.41152	15.9%



a) Figure S3 Related to Figure 4. a) Sequence alignment of Siglec-10 CRD (aa 17-144) respect to the CD33-like Siglecs templates, namely Siglec-3, Siglec-5, Siglec-7, Siglec-8. The residues responsible for the binding of sialylated glycans are highlighted in green. **b)** Superimposition of the best Siglec-10 homology-based model (orange) and the X-ray crystal structures from other Siglec members. The residues of Siglec-10 that interacted with the sialic acid moiety which are reported to be crucial for the interaction with sialylated glycans are shown as sticks. The Arg in position 119 in Siglec-10 was indeed conserved in all the Siglecs, along with the aromatic residues in position 21 and 129 (Phe and Tyr in Siglec-10).

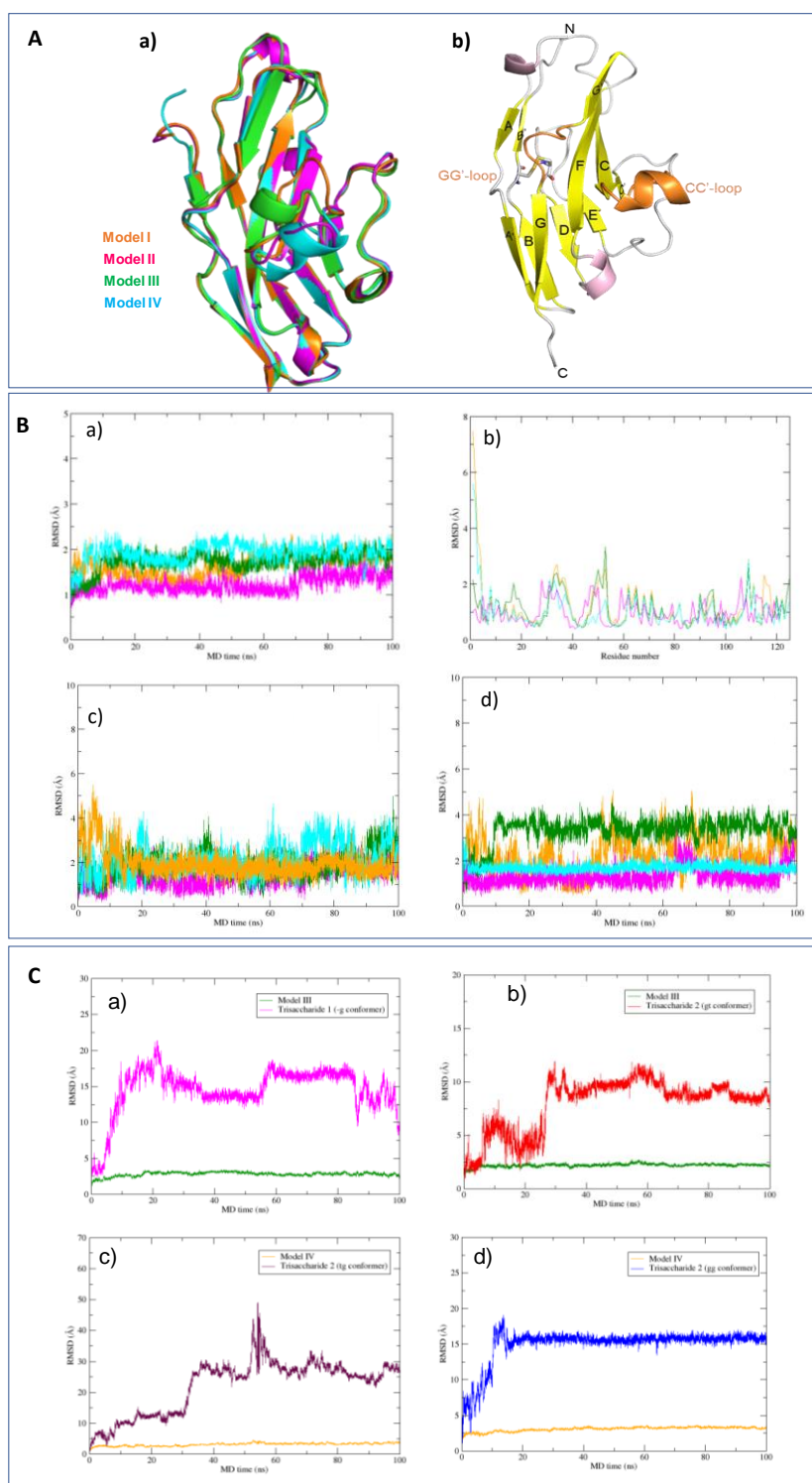


Figure S4 Related to Figure 4. A. Structural features of Siglec-10 selected model. a) Superimposition of the selected homology models. Relevant differences can be observed only in one of the variable Siglec region that is the CC' loop. b) Structure of the best Siglec-10 model obtained by means of Swiss Model by using Siglec-7 as template. The secondary structure elements are indicated according to Siglecs V-set domain nomenclature. **B Analysis from the MD simulations of the best homology models.** a) Root mean square deviation (RMSD) of the backbone of Model I (orange), Model II (purple), Model III (green), Model IV (cyan) b) Root mean square fluctuation (RMSF) of the backbone of the different models of Siglec-10 c-d) RMSD of the backbone of the important loop regions (CC'-loop) on the left and (GG' -loop) on the right. **C MD simulation trajectory analysis of different Siglec-10/ligand complexes from homology modeling and docking studies.** a) MD simulation analysis of model III/1 complex (-g conformer). b) MD simulation analysis of model III/2 complex (gt conformer) c) MD simulation analysis of model IV/2 (gg conformer) complex. d) MD simulation analysis of model IV/2 (tg conformer) complex. The ligand in complexes a) and c) showed low stability during the MD simulation, further confirmation that could be excluded as model of interaction.

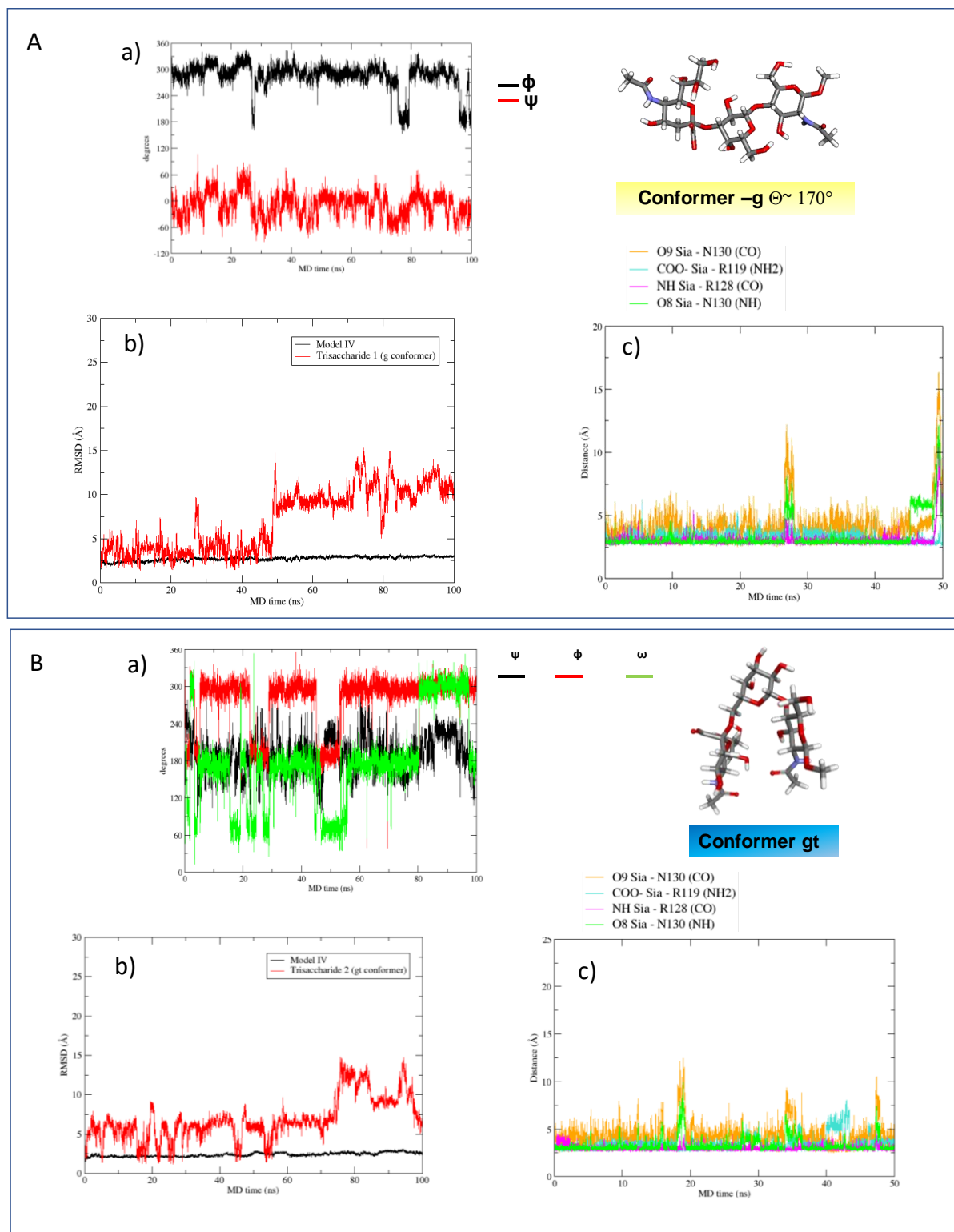
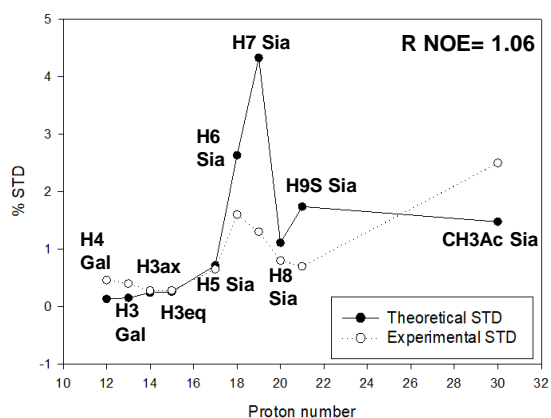
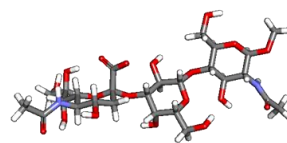


Figure S5 Related to Figure 4. MD simulation trajectory analysis of ligands/model IV complex (upper panel) and trisaccharide **2**/model IV complex (lower panel). **Upper panel A:** a) Dihedrals ϕ , ψ of Neu5Ac- α -(2,3)-Gal linkage fluctuation in MD simulation of the bound state (model IV); b) RMSD plot of trisaccharide **1** (-g conformer) in complex with Siglec-10 CRD (model IV). c) Representative intermolecular distances between Neu5Ac unit and the Siglec-10 residues calculated along the simulation. According to this figure several hydrogen bonds in the complex can be expected. **Lower panel B:** a) Dihedrals ϕ , ψ , ω of Neu5Ac- α -(2,6)-Gal linkage fluctuation in MD simulation of the bound state; b) RMSD plot of trisaccharide **2** (gt conformer) in complex with Siglec-10 CRD (model IV). c) Representative intermolecular distances between Neu5Ac unit and the Siglec-10 residues calculated along the simulation. According to this figure several hydrogen bonds in the complex can be expected.

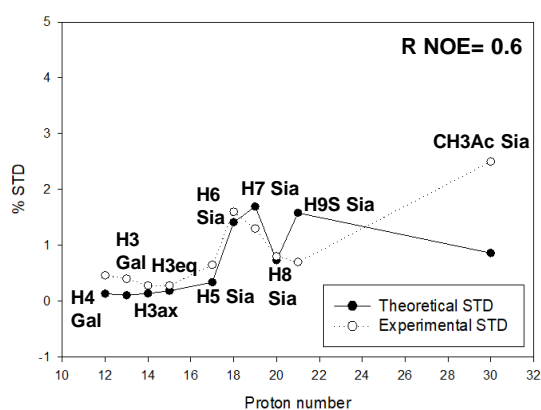
A



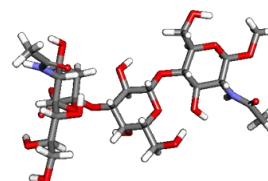
model IV/1 complex



Conformer g

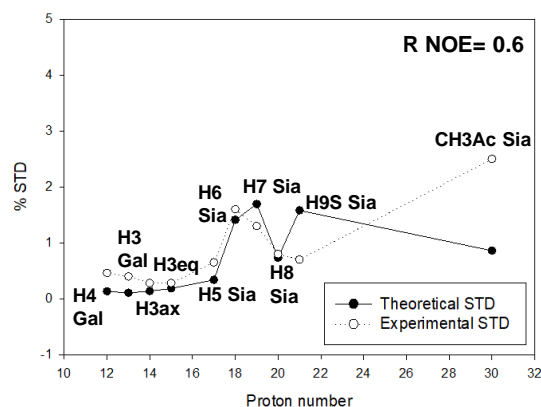


model IV/1 complex

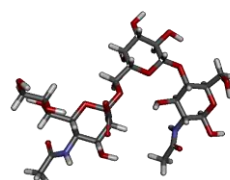


Conformer t

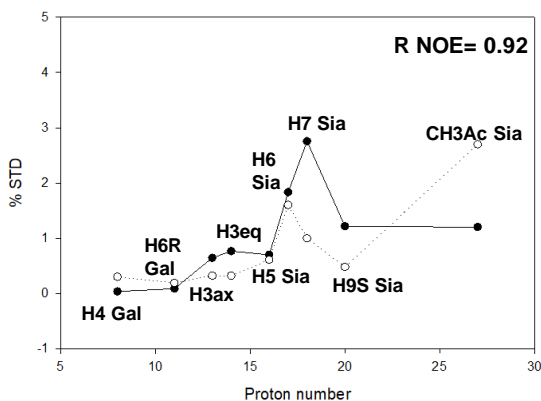
B



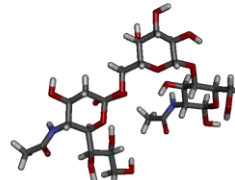
model IV/2 complex



Conformer gg



model IV/2 complex



Conformer tg

Figure S6 Related to Figure 4: CORCEMA-ST analysis of different models of binding generated by the modelling studies with 1 (upper panel) and 2 (lower panel).

For all the reported models, the profile of the experimental STD data was not similar from that of the theoretical STD effects. The predicted STD values were very different from the experimental ones for several protons suggesting that these models should be excluded.

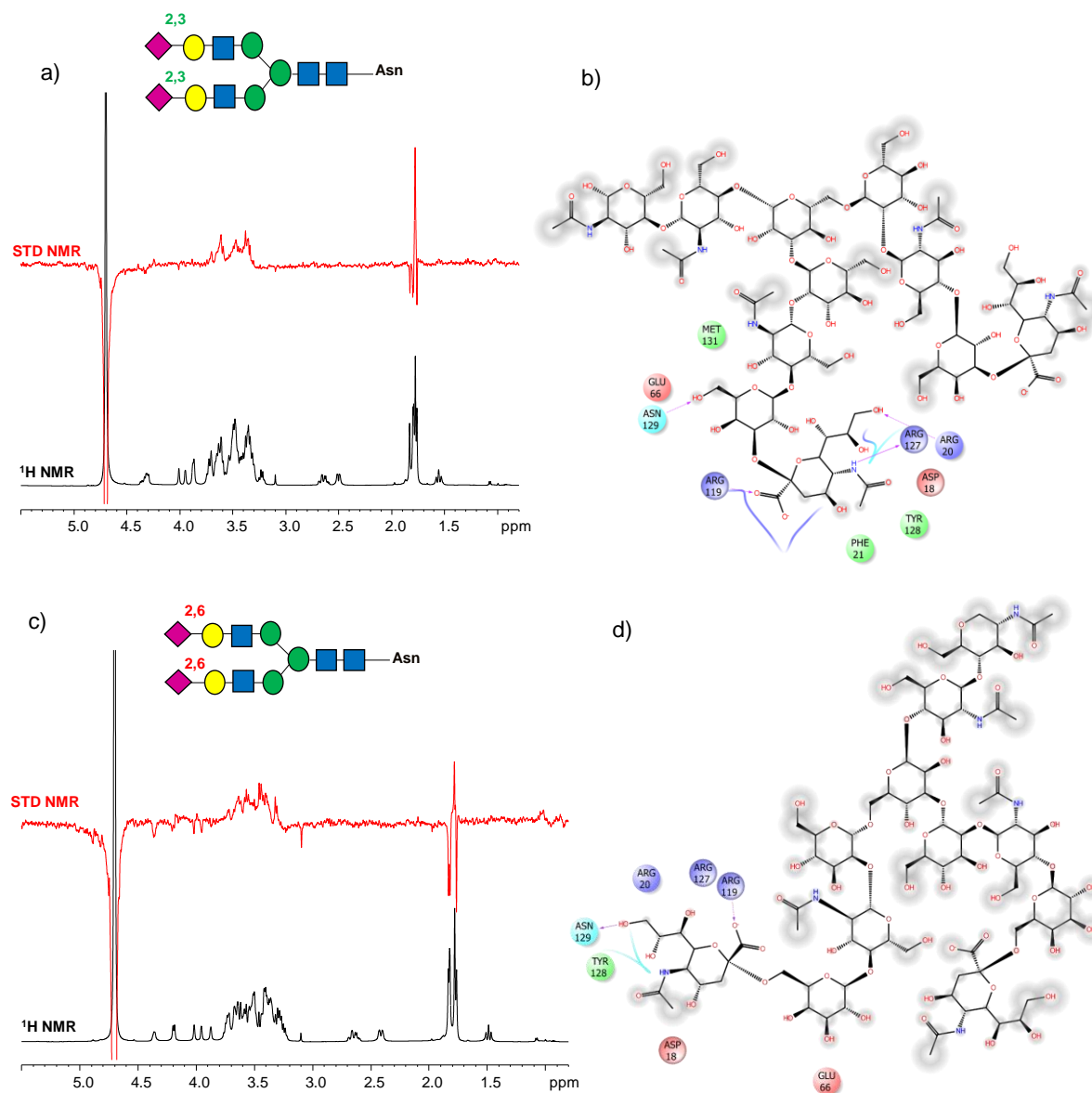


Figure S7 Related to Figure 5. Analysis of the interactions between longer N-glycans and Siglec-10. a) STD NMR interaction between Siglec-10 and undecasaccharide 3. b) Binding pose predicted by the docking of the undecasaccharide 3 bound form into the binding pocket of Siglec-10 model. c) STD NMR analysis of the interaction between Siglec-10 and undecasaccharide 4; d) Binding pose predicted by the docking of the undecasaccharide 4 bound form into the binding pocket of Siglec-10 model.

Transparent Methods

NMR analysis: All NMR experiments were recorded on a Bruker 600-MHz DRX equipped with a cryo probe and all the samples were dissolved in 50mM phosphate buffer, pH 7.4. Data acquisition and processing were performed with TOPSPIN 3.2 software.

Data sets of 4096x512 ($t_1 \times t_2$) points and zero-filled in the F1 dimension up to 4096x2048 points were used when homonuclear spectra were measured. Before Fourier transformation, a cosine-bell function was used. Heteronuclear single quantum coherence (HSQC) experiments were recorded in the ^1H -detected mode by single quantum coherence with proton decoupling in the ^{13}C domain, with data sets of 2048x256 points. The phase-sensitive mode by the method of States et al. was used in all the experiments (States et al., 1982).

Tr-NOESY analysis: For ligand **1**, one-dimensional selective NOE experiments were acquired at mixing times ranging from 50-800ms, at 283K, following the selective excitation of proton H3 of Gal residue.

For ligand **2**, homonuclear 2D ^1H - ^1H NOESY experiments were performed by using data sets of 4096x512 points and mixing times between 50 and 800 ms.

The protein: ligand ratio varied from 1:10 to 1:50. Proton – proton cross relaxation rates were measured integrating the NOE signals of interest. The raw data were normalized against the corresponding diagonal/selectively irradiated peak that was fitted to an exponential decay function and extrapolated back to an intensity of 100% at zero mixing time. Once normalized, the intensities were plotted against the mixing times and the build-up curves were fitted to an exponential equation of the form $f(x) = A \cdot (1 - \exp(-b \cdot x)) \cdot (\exp(-c \cdot t))$, with A, b and c being adjustable parameters. The initial slope at 0 mixing time (σ) of each curve was used to calculate the experimental distances (r_{ij}) by employing the isolated spin pair approximation. The intra-residue NOE contacts used as reference for distance determinations were: H1-H3 of galactose moiety for ligand **1** and H1-H5 of the N-acetylglucosamine residue A was for ligand **2**. The experimental error in the calculated proton-proton distances is estimated to be less than $\pm 10\%$.

STD NMR analysis: Control experiments with the ligands in the absence of the protein were performed to optimize STD NMR parameters.

STD NMR experiments were recorded with 32 k data points and zero-filled up to 64 k data points prior to processing. The protein/ligand molar ratio spanned from 1:50 to 1:200. 40 Gauss pulses with a length of 50 ms were used to selectively irradiate the protein resonances by using the on-resonance pulse at 7.5 ppm, while the off-resonance pulse frequency was at 40 ppm. The unwanted broad signals of the receptor were reduced by applying a 20 ms spin lock pulse.

The fractional STD effects were calculated by use of $(I_0 - I_{\text{sat}})/I_0$, where I_{sat} is the intensity of the signal in the STD NMR spectrum and I_0 is the peak intensity of an unsaturated reference spectrum (off-resonance). The construction of STD build up curves was performed by fitting the saturation time data to a monoexponential equation of the form: $\text{STD} = \text{STD}_{\text{max}}[1 - \exp(-k_{\text{sat}}t)]$, where STD is for the STD signal intensity of a given proton at a saturation time t, STD_{max} stands for the asymptotic maximum of the curve, and k_{sat} represents the observed saturation rate constant measuring the speed of STD build-up. The value of STD_{fit} was derived by the slope of the STD build-up curve at a saturation time of 0. Once calculated both STD_{fit} and K_{sat} values, all the intensities of different protons ligand were normalized to the largest STD_{fit} , giving $\text{STD}_{\text{epitopes fit}}$.

CORCEMA-ST protocol: CORCEMA-ST protocol has been used as previously described (Wen et al., 2005; Bhunia et al., 2004). The pdb coordinates of complexes were generated from Docking and MD analyses and the conformation of the ligand was considered to be the same in both free and bound state. The input variables, including the concentration of the ligand and the protein, were deduced experimentally. The dissociation constants (KD) were set in the range of 10^{-3} – 10^{-6} M and modified further to get the best fit. A binding site cut-off of 8 Å was employed. After the computation of R matrix and the calculation of spectral densities, the predicted fractional intensity changes for each ligand protons were compared to the experimental STD effects by using a NOE R factor, a normalized root-means square deviation (RMSD) value.

Preparation of glycans: Ligand **1** and **2** were purchased from Tokyo Chemical Industry Co., Ltd. (Product Number: N0949, N0950). Ligand **3** and **4** were provided from GlyTech, Inc (Kajihara et al., 2004).

Protein expression and purification: A cDNA encoding a recombinant form of human Siglec-10 containing Chinese Hamster Ovary cell line stably expressing the 5 3 N-terminal Ig-like domains of human Siglec-10 CD22 fused to the Fc region of humanmouse IgG1G2b was prepared as described previously (Munday et al., 2001), cloned into the pEE14 glutamine synthetase expression system (Bebbington et al., 1992) and transfected into kindly provided by Dr Ajit Varki (University of California San Diego) Chinese Hamster Ovary (CHO) cells. A CHO cell clone stably expressing Siglec-10-Fc was expanded in glutamine-free Glasgow's modification of Eagles Medium (GMEM) containing 200 micromolar L-methionine sulfoximine and 10% FCS. Once cells were ~90% confluent, the medium was switched to GMEM + 0.5% Hyclone Fetalclone II (low IgG) and cells left for up to 7 days. Three Cells were expanded in flasks in Hams F10 + 10% FCS and when confluent the medium was changed to Hams F10 + 0.5% FCS and cells incubated for up to 7 days. 4 litres of supernatants were pooled, 0.2 μM filtered and passed over a 1.0 ml protein A-Sepharose column. Following washing with 100 ml PBS, the Siglec-10-Fc protein was eluted with 0.1 M glycine pH 3.0 and neutralised with one tenth volume of 1.0 M Tris buffer pH 8.0. Fractions containing > 0.1 mg/ml protein were pooled, dialysed

against PBS and stored at -80°C. Purity of the protein was assessed by SDS-PAGE under reducing and non-reducing conditions followed by Coomassie Blue staining.

Homology modeling. The carbohydrate recognition domain of human Siglec-10 was considered for computational 3D structure calculation by homology modeling. The sequence encoding for the human Siglec-10 (accession No. Q96LC7.3) was extracted from NCBI (<http://www.ncbi.nlm.nih.gov>). The carbohydrate recognition domain was identified by submitting the query sequence to CD-search (Marchler-Bauer et al., 2004; Marchler-Bauer et al., 2017). As a result, the region between the residues 23-140 of the protein query was found to match ('specific hit') to the Immunoglobulin (Ig) domain at the N terminus of Siglec (accession No. cd05712). The previous interval was then submitted to the BLAST search (Altschul et al., 1990; Gish et al., 1993), using protein data bank as database to identify potential templates for the homology modeling of Siglec-10. Siglec-8 (PDB-ID:2N7A), Siglec-7 (PDB-ID:2HRL), Siglec-5 (PDB-ID:2ZG2) and Siglec-3 (PDB-ID:5IHB) were selected as the templates because of the highest identity (47-56%) and query cover (85-99%). For each template, four homology models were generated using four different servers: SWISS-MODEL (Arnold et al., 2006; Biasini et al., 2014), I-TASSER (Roy et al., 2010; Yang et al., 2015), PHYRE2 (Kelley et al., 2015), and RAPTOR-X (Kallberg et al., 2012). (Figures S3-S4A). To select the best models, detailed analysis was performed, focusing on the orientation of the binding site residues and the flanking loops. The models derived from I-TASSER, and PHYRE-2 were discarded because the structures did not have coherence with the experimental 3D structures of the homologous proteins. Thus, computational studies were performed using the 3D structures obtained from SWISS-MODEL and RAPTOR-X, which were submitted to molecular dynamics (MD) simulations of 100 ns to optimize the geometry. To assess the quality of the homology models, the results were compared to the MD simulation (100 ns) performed with the crystal structure of Siglec-3 (PDB-ID: 5IHB), which was carried out using the same parameters. The macromolecule was chosen as reference because of the similarity to Siglec-10, in terms of binding specificity and affinity toward sialylated glycans (Varki et al., 2006).

Each of the computed models displayed high similarity in terms of 2D and 3D structure among them and also in comparison to the corresponding templates, showing a nearly invariant canonical V-set Ig-domain fold. As can be expected by the sequence alignment (Figure S3A), the superimposition to other Siglec crystal structures revealed some diversity concerning the amino acids constituting the binding site (Figure S3B), pointing to a putative role in the specificity of recognition of sialylated glycans. Nevertheless, the selected models exhibited different shape in the CC' loop, probably due to the lower sequence similarity in that region with respect to the structural templates (Figure S3), leading to difficulties in the loop conformation prediction. For this reason, the mobility of the CC' and GG'-loop of the selected models has been investigated during the dynamics (Figures S4A and B). From the MD simulation results was observed that the models obtained using Siglec-3 and Siglec-7 as templates (both the Swiss-Model and the Raptor-X models) displayed higher stability along the simulations and therefore they were used for the docking calculations (Figure S4).

Molecular dynamics simulations. By means of MD simulations as implemented in AMBER 14 suite of programs it was possible to investigate the ligands behavior in solution, to assess the stability of the homology models, the mobility of the loops and the stability of the docking poses. Missing hydrogen atoms were added, and protonation state of ionisable groups was computed by using Maestro Protein Preparation Wizard (Schrodinger et al., 2012; Schrodinger et al., 2012; Schrodinger et al., 2012). Atom types and charges were assigned according to AMBER ff14SB, for the homology models and GLYCAM-06j-1 force field for the ligands. To simulate the aqueous environment the homology models and ligands were hydrated by using octahedral boxes containing explicit TIP3P water molecules extending 10 Å away from any atom. Counter ions were added to neutralize the system by using the Leap module. Minimization was performed using Sander and MD simulations were carried out using the CUDA, which are distributed within the AMBER 14 package. The smooth particle mesh Ewald method was used to represent the electrostatic attractions in the system while each simulation was under periodic boundary conditions, and the grid spacing was 1 Å. Initial annealing of the system occurred steadily and lightly from 100 °K to 300 °K over 25 ps. Temperature was kept constant at 300 °K during 50 ps with progressive energy minimizations and a solute restraint. The restraints were gradually released by the solute, which was closely followed by a 20 ps heating period which went from 100° K to 300° K, once completed the restraints were removed. The systems then advanced in an isothermal-isobaric ensemble. Long-range electrostatic interactions were accounted for by means of the particle mesh Ewald approach applying periodic boundary conditions. Each of the simulations lasted 100 ns. Coordinates were archived in order to acquire 1000 structures of the progression of the dynamics. Trajectories were analysed with the ptraj module included in the AMBER14 software package and visualized with VMD molecular visualization program (Roe et al., 2013).

Ligand-protein docking calculations.

Preparation of the macromolecules. 3D coordinates of Siglec-10 homology models were used for docking purposes. In addition, the docking calculations were performed by using the 1 ns average structure extracted by MD simulations of the computed Siglec 10 models. Then, each structure was submitted to 100 000 steps of steepest descent minimization with MacroModel and optimized with AMBER force field, before being used

for docking calculations. Quality of the minimized models was verified using the procheck online resource (Laskowski et al., 1993) that displayed Ramachandran plot of the structures.

Building of ligands. The 3D coordinates of [Neu5Ac- α -(2,3)-Gal- β -(1,4)-GlcNAc- β] and [Neu5Ac- α -(2,6)-Gal- β -(1,4)-GlcNAc- β] were built with the help of xLEAP using the GLYCAM_06j-1 force field parameters. It is worth to note that the relative orientation of saccharide units is expressed in terms of the glycosidic linkage torsion angles ϕ and ψ , for α (2-3) linkage and ϕ and ψ and ω for α (2-6) linkage. Therefore to investigate the conformational behavior of the ligands, the three conformers (*g*-, *g*+, *t*) of Neu5Ac- α -(2,3)-Gal- β -(1,4) glycosidic linkage and the three rotamers (*gt*, *gg*, *tg*) of Neu5Ac- α -(2,6)-Gal- β were considered for docking calculations. For each conformer, the bonds were parametrised and the geometries were optimized by 100000 step of steepest descent minimization with OPLS3 force field by using Macro Model. The resulting structures were submitted to 100 ns MD simulations to assess the stability. Ligands were prepared for docking calculations using AutoDockTools, setting all rotatable bonds free to move, except for the glycosidic linkages, during the docking calculations.

Docking calculations. Docking calculations of all compounds were performed by using AutoDock 4.2.2. (Morris et al., 2009). Analysis of the docking poses was performed with AutoDockTools. The docking protocol was validated by carrying out the docking of Siglec-3 crystallographic structure in complex with α (2-3) Neu5Ac-Gal ligand (PDB-ID: 5J06). The 3D structure of α (2-3) Sia-Gal was extracted from the crystallographic structure of Siglec-3. The grid point spacing was set at 0.375 Å, and a hexahedral box was built with x, y, z dimensions: 64 Å, 46 Å, 56 Å centered in the centroid position among residues Arg119, Arg127 and Asn129 for Siglec-10 models; Arg121, Lys128 and Ser130 for Siglec-3. A total of 200 runs using Lamarckian Genetic algorithm was performed, with a population size of 100, and 250000 energy evaluations.

Molecular modelling and CORCEMA of 1 and 2 in complex with Siglec-10

Docking calculations of trisaccharide **1** in its three possible conformations within the selected modeled structures (model I, model II, model III, model IV) were performed.

The docking results showed that no reasonable binding pose was predicted for the models I and II in none of the possible ligand conformations. Additionally, the docking calculations performed using 1 ns average structure calculated by MD simulations didn't led to rational representation of the interacting surface. The models III and IV, instead, resulted in binding poses consistent to the experimental STD and NOE data. Also, in these complexes, the polar contacts established between the ligand and the receptor were in agreement with those observed for experimental binding poses of homologous proteins with similar ligands (Zhuravleva et al., 2008; Attrill et al., 2006). The selected protein-ligand complexes were submitted to 100 ns MD simulations. In detail, by comparing the 3D complexes obtained from the docking of trisaccharide **1** in the *g*-conformation with the two foremost structural models, it was noted that the higher stability along the simulation was achieved by using model IV rather than model III (Figures S4C and S5A). As a consequence, the receptor model III was excluded from the CORCEMA ST characterization.

Then, the 3D complexes obtained from model IV homology, which presented different trisaccharide **1** geometry (*g* and *t* conformers), were submitted to CORCEMA-ST calculation giving a R NOE value significantly higher than model IV/**1** complex (Figure S6A). This outcome permitted to further assess the selection of the *g*-conformer upon binding.

Docking calculations of ligand **2** were performed in the selected Siglec-10 modeled structures. From the docking studies three reliable binding poses of **2** were predicted using model IV as receptor. The binding poses differed essentially for the value of the ω angle around the Neu5Ac α (2,6) Gal glycosidic linkage: In the three selected binding poses, one pose of the ligand in the *gt* conformation ($\omega = 60^\circ$), one pose in the *tg* conformation ($\omega = 180^\circ$) and the last in the *gg* conformation ($\omega = -60^\circ$), the ligand was placed with good orientation into the binding site of Siglec-10. Therefore, the obtained complexes were submitted to 100 ns MD simulations to assess the stability (Figures S4C and S5B). From the simulation results it was possible to observe significantly higher stability for the *gt* conformer complex. Indeed, the trisaccharide **2** binding mode within the Siglec-10 model IV structure showed stability until 70 ns and maintained all the described polar interactions without restrictions. It is worth to note that the CORCEMA-ST allowed to finally select the best complex, as for the model IV/ **2** *tg* and *gg* conformers, the predicted STD values were very different from the experimental ones for many protons (Figure S6B).

Steady-state fluorescence spectroscopy

Steady-state fluorescence spectra have been acquired on a Fluoromax-4 spectrofluorometer (Horiba, Edison, NJ, USA) at the fixed temperature of 5°C. Emission spectra were recorded in the emission range of 290–500 nm upon excitation at 280 nm. The slit widths were fixed at 4 nm for the excitation and 10 nm for the emission wavelength. A quartz cuvette with a path length of 1 cm and a chamber volume of 1 mL was used under constant stirring. A 0.9 mL volume of protein solution at fixed concentration of 0.2 μ M was titrated by adding small aliquots (1–20 μ L of a ligand stock solution of 500 μ M) of trisaccharide **1** and **2**. The PBS buffer at pH 7.4 was used for all solutions. The optical density of the solution at the excitation wavelength was kept less than 0.05. Small errors due to dilution upon ligand additions were neglected. The protein fluorescence was found to quench in the presence of the ligands. The binding curve was obtained by plotting F/F_0 values versus

ligand concentration. Data analysis were performed using the following equation (Ribeiro et al., 2008; Oliva et al. 2019).

$$\phi_{FY} \frac{F}{F_0} = 1 + \left(\frac{\phi_{FY}}{\phi_F} - 1 \right) X_{FY}$$

where F_0 and F are the integrated areas of the emission spectra in the absence and the presence of ligands, and ϕ_{FY} and ϕ_F are the quantum yield of the fluorophores in the complex and in the protein, respectively. The term X_{FY} can be written as:

$$X_{FY} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where $a = [F]_t K_b$, $b = 1 + [Y]_t K_b$ and $c = [Y]_t K_b$; $[F]_t$ and $[Y]_t$ represent the total concentration of protein and ligand, respectively.

ELISA: a solution of 50 μ L of synthetic polyacrylamide (PAA) polymers functionalized with Neu5Ac α 2-3Gal β 1-4Glc β -NHCOCH₂NH₂ or Neu5Ac α 2-6Gal β 1-4Glc β -OCH₂CH₂NH₂ (purchased from GlycoNZ, Auckland, New Zealand, carrying the same moles of carbohydrate) at 200 μ g/mL, or a solution of 50 μ L of synthetic glycopeptides containing sialylated *N*-glycans at 1mg/mL, in PBS (10mM, pH=7.4), were used to coat the Nunc MaxiSorp plate 2h at room temperature. After discarding and washing with PBS (1x150 μ L), the wells were blocked with 100 μ L of 1% BSA (Sigma-Aldrich, lyophilized powder, \geq 96%, agarose gel electrophoresis) in PBS at room temperature for 30 min. The blocking solution was discarded and 50 μ L of Fc-Siglec-10 at 1 μ g/mL (Recombinant Human Siglec-10 Fc Chimera Protein from R&D) in assay buffer (0.5% BSA) were added to the wells. After 1h at 37°C, the wells were washed with PBS (2x150 μ L) and then 100 μ L of anti-human horseradish peroxidase (0.8 μ g/mL, life technologies, Novex® Goat anti-Human IgG-HRP) were added. After 30 min at room temperature, the wells were washed with PBS (2x150 μ L). Finally, 100 μ L of substrate solution (3,3',5,5'-Tetramethylbenzidine, TMB, in citric/acetate buffer, pH=4, and H₂O₂) were added and after 5 min incubation at room temperature the reaction was stopped with 50 μ L of H₂SO₄ (0.8 M) and the optical density (OD) was measured at 450 nm in an ELISA reader. The experiment was performed in duplicate and data were normalized over the signal at 450 nm from the BSA-containing wells.

References

- States, D. J., Haberkorn, R. A., Ruben, D. A. (1982) A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants*. *J. Magn. Reson.*, 48, 286–292.
- Wen, X., Yuan, Y., Kuntz, D. A., Rose, D. R. Pinto, B. M. (2005) A Combined STD-NMR/Molecular Modeling Protocol for Predicting the Binding Modes of the Glycosidase Inhibitors Kifunensine and Salacinol to Golgi α -Mannosidase II. *Biochemistry*, 44, 6729-6737.
- Bhunia, A., Jayalakshmi, V., Benie, A. J., Schuster, O., Kelm, S., Krishna, N. R., Peters, T. (2004) Saturation transfer difference NMR and computational modeling of a sialoadhesin-sialyl lactose complex. *Carbohydr. Res.* 339, 259-267.
- Kajihara, Y., Suzuki, Y., Yamamoto, N., Sasaki, K., Sakakibara, T., Juneja, L.R. (2004) Prompt chemoenzymatic synthesis of diverse complex-type oligosaccharides and its application to the solid-phase synthesis of a glycopeptide with Asn-linked sialyl-undeca- and asialo-nonasaccharides. *Chem. Eur. J.*, 10, 971-985.
- Munday, J., et al. (2001) Identification, characterization and leucocyte expression of Siglec-10, a novel human sialic acid-binding receptor. *Biochem J.* 355, 489-497.
- Bebbington, C. R., et al. (1992). "High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker." *Biotechnology (N Y)* 10(2), 169-175.
- Marchler-Bauer and A. Bryant S. H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.*, 32, 327-31.
- Marchler-Bauer, A., et al. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.*, 4, 45, D200-D203.
- Altschul, S.F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215, 403-410.
- Gish, W. and States D. J. (1993) Identification of protein coding regions by database similarity search. *Nature Genet.*, 3, 266-272. DOI: 10.1038/ng0393-266.
- Arnold, K., Bordoli, L., Kopp, J., Schwede T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. *Bioinformatics*, 22, 195-201.
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, Bordoli, M. L., Schwede, T. (2014) SWISS-MODEL: modeling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.*, 42, W252-8.
- Roy, A., Kucukural, A., Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc.*, 5, 725-38.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods*, 12, 7-8.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, 10, 845-58.

Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., Xu, J. (2012) Template-based protein structure modeling using the RaptorX web server. *Nat Protoc*, 7, 1511-22.

Varki, A., Angata T. (2006) Siglecs—the major subfamily of I-type lectins, *Glycobiology*, 16, 1, 1R–27R.

Schrodinger. 2012. Epik version 2.3, Schrodinger, LLC, New York, NY.

Schrodinger. 2012. Impact version 5.8, Schrodinger, LLC, New York, NY, 2012.

Schrodinger. 2012. Prime version 3.1, Schrodinger, LLC, New York, NY, 2012.

Roe, D.R., Cheatham, T. E. (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data *J. Chem. Theory Comput.* 9, 3084-3095.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993) PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.*, 26, 283-291.

Morris, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., Olson, A. J. (2009) AutoDock treats the ligand as a flexible unit and the protein as a rigid unit. *J Comput Chem.*, 30, 16, 2785–2791.

Zhuravleva, M. A., Trandem, K., Sun, P. D. (2008) Structural Implications of Siglec-5-Mediated Sialoglycan Recognition, *Journal of Molecular Biology* 375, 437-447.

Attrill H. et al. (2006) The structure of siglec-7 in complex with sialosides: leads for rational structure-based inhibitor design *Biochemical Journal.*, 397 (2) 271-278.

Ribeiro, M. M. B.; Franquelim, H. G.; Castanho, M. A. R. B.; Veiga, A. S. (2008) Molecular Interaction Studies of Peptides Using Steady-State Fluorescence Intensity. Static (De)Quenching Revisited *J. Pept. Sci.* 14 (4), 401– 406.

Oliva, R., Battista, F., Cozzolino, S. et al. (2019) Encapsulating properties of sulfobutylether- β -cyclodextrin toward a thrombin-derived antimicrobial peptide. *J Therm Anal Calorim* 138, 3249–3256.